

26643 Overheidsbrede visie op generatieve artificiële intelligentie (AI)

1125 Informatie- en communicatietechnologie (ICT)

II Antwoord / Reactie van de bewindspersoon

I Vragen en opmerkingen vanuit de fracties

Vragen en opmerkingen van de leden van de GroenLinks-PvdA-fractie en antwoord

De leden van de GroenLinks-PvdA-fractie hebben met interesse kennisgenomen van de Overheidsbrede visie generatieve AI. Deze leden erkennen dat generatieve AI, mits goed ingezet, een belangrijke maatschappelijke rol kan vervullen. Echter zijn de mogelijke risico's dusdanig groot dat de leden kritisch blijven op het uitrollen en omarmen van deze te weinig gereguleerde technologie. Daarover hebben zij opmerkingen en vragen.

De leden van de GroenLinks-PvdA-fractie zullen hun opmerkingen en vragen per hoofdstuk benoemen, te beginnen met hoofdstuk 1. Hierin wordt verwezen naar het wettelijk kader dat de AI-verordening schept. Deze leden vragen op welk gebied er aanvullend nationaal beleid nodig is om volledig te handelen naar de waarden die Nederland nastreeft in haar digitalisering. (1)

Als het gaat om het reguleren van AI binnen de EU en in Nederland, vormt de AI-verordening het juridisch fundament. Dit in aanvulling op onder meer de Algemene verordening gegevensbescherming (AVG), de Algemene wet bestuursrecht (Awb) en de Algemene wet gelijke behandeling. De AI-verordening heeft als doel ervoor te zorgen dat AI-systemen die op de EU-markt worden gebruikt, veilig zijn of de fundamentele rechten en waarden van de EU respecteren, ongeacht of die AI-systemen binnen of buiten de Unie zijn ontwikkeld.

Daarnaast heeft het kabinet gewerkt aan verschillende (niet-regulerende) beleidsacties op nationaal niveau die moeten bijdragen aan het borgen van onze publieke waarden binnen de (digitale) samenleving. Dit beleid is onderdeel van onder meer de Werkagenda Waardengedreven Digitaliseren en de Strategie Digitale Economie. Voorbeelden hiervan zijn bijvoorbeeld het algoritmeregister, de ELSA-labs¹ en het programma AiNed. Daarbij is er de afgelopen jaren nog meer specifiek (nationaal) beleid ontwikkeld met het oog op het beschermen van grondrechten en het stimuleren van innovatie; dit beleid wordt nader toegelicht in hoofdstuk 4a van de overheidsbrede visie op generatieve AI (onder het kopje 'AI-beleid').²

De leden van de GroenLinks-PvdA-fractie hebben opmerkingen en vragen over hoofdstuk 2. Zij vinden het risico van 'hallucineren', waarbij een AI-model een verkeerd antwoord geeft, groot. Is het mogelijk om, bij AI-gegenereerde antwoorden die onzeker of niet geverifieerd zijn, een indicatie van twijfel toe te voegen? Welke waarborgen zijn er of worden ingebouwd zodat AI-modellen niet naar onbetrouwbare websites verwijzen en deze informatie uit de gebruikte trainingsdata wordt geweerd? (2)

Het kabinet erkent de zorgen van de leden van de GroenLinks-PvdA-fractie. De term hallucinatie verwijst naar door large language models (LLM's) gegenereerde content die door het model wordt gepresenteerd als feitelijk juist maar dit niet is. Dit probleem is mogelijk één van de grootste uitdagingen wanneer het gaat om veilige implementatie van LLM's in systemen die impact hebben op de samenleving. Hallucinaties zijn in het onderzoeksveld 'natural language processing' een bekend probleem en een belangrijk aandachtspunt. Zij hebben verschillende oorzaken die voortvloeien uit verschillende stappen in de ontwikkelings- en finetuningsfase van LLM's. Er bestaat een breed scala aan manieren van aanpak om hallucinaties terug te dringen, waarbij deze verschillende aanpakken aansluiten op de verschillende oorzaken. Zo zijn er technieken gericht op het opschonen van de input-data, of het toevoegen van externe databases waarvan de inhoud kan

¹ Dit zijn labs waarin door wetenschappers, ondernemers en publieke instellingen onderzoek wordt gedaan naar de ethische, juridische en sociale aspecten van AI.

² <https://open.overheid.nl/documenten/9aa7b64a-be51-4e6a-ad34-26050b8a67ef/file>

worden gevalideerd.³ Echter is het op dit moment middels bestaande technieken (nog) niet mogelijk hallucinatie volledig te voorkomen. Hallucinaties kunnen zich blijven voordoen, zelfs na het toepassen van diverse mitigatiestrategieën. Het kabinet heeft daarom in internationaal verband (bijvoorbeeld via de OESO en de Verenigde Naties) gepleit voor veiligheids- en interpretatieregels voor generatieve AI-modellen en stimuleert onderzoek naar verantwoorde en transparante AI-modellen. Bovendien is, zoals ook opgenomen in de overheidsbrede visie op generatieve AI, door het kabinet ingezet op het versterken van digitale vaardigheden en digitaal bewustzijn van mensen in Nederland zodat zij bewust, kritisch en actief kunnen omgaan met AI-toepassingen en de risico's die daarmee gepaard gaan, zoals het risico op hallucinaties.

Om de ontwikkeling van Nederlandse taalmodellen te stimuleren is daarbij recent 13,5 mln. euro vanuit EZK/RVO toegekend voor de ontwikkeling van GPT-NL, een NL-taalmodel gebaseerd op publieke waarden en voor publieke doelen. Hiermee wordt een open en transparant Nederlands taalmodel ontwikkeld dat oog heeft voor vraagstukken rondom betrouwbaarheid en transparantie.

De AI-verordening bevat daarnaast eisen aan AI-modellen die in staat zijn om veel verschillende taken uit te voeren en in veel verschillende AI-systemen kunnen worden geïntegreerd. Hieraan wordt gerefereerd als 'AI-modellen voor algemene doeleinden'.⁴ Aanbieders van deze modellen zijn verplicht om informatie over de werking van het model te delen met ontwikkelaars die op dit model voortbouwen. Ook moet duidelijk worden hoe het auteursrecht wordt gerespecteerd en moet bij alle door AI gegenereerde content duidelijk worden gemaakt dat het dit het geval is. Deze eisen dragen er aan bij dat beter inzichtelijk wordt welke content tot stand is gekomen met AI en mogelijk onzeker, fout of ongeverifieerd is.

Er gelden aanvullende eisen als het AI-model voor algemene doeleinden volgens de AI-verordening systeemrisico's met zich meebrengt. De bedoeling is dat de AI-verordening hiermee de krachtigste AI-modellen op de Europese markt reguleert. De aanbieder van een dergelijke AI-model ook die systeemrisico's in kaart brengen en mitigeren. Dit geldt voor modellen met veel capaciteiten, bijvoorbeeld als gevolg van de grote rekenkracht die wordt gebruikt om het model te trainen. Het consequent doorverwijzen naar onbetrouwbare website of het geven van onjuiste informatie zou mogelijk als systeemrisico kunnen worden gezien. Of dit daadwerkelijk het geval is, is ook afhankelijk van de nog op te stellen codes of practice en richtsnoeren. Deze worden opgesteld door het Europese AI-bureau, het bureau binnen de Europese Commissie dat verantwoordelijk wordt voor de handhaving van eisen aan AI-modellen voor algemene doeleinden.

Ook lezen de leden over het oprichten van een AI-adviesraad. Zij vinden het wenselijk dat deze als kenniscentrum dient. Echter zijn de leden van deze fractie van mening dat de discussie rondom waarden en ethiek bij het toepassen van AI in de Tweede Kamer thuishoort. Welke bevoegdheden heeft de adviesraad en over welke thema's zullen zij adviezen uitbrengen? (3)

Het is van groot belang om de discussie rondom waarden en ethiek bij de toepassing van AI in uw Kamer te voeren. De afgelopen tijd hebben er vanuit de ministeries van Binnenlandse Zaken en Koninkrijksrelaties, Economische Zaken en Klimaat en Onderwijs, Cultuur en Wetenschap vele gesprekken plaatsgevonden om te verkennen wat de mogelijkheden zijn voor een dergelijke adviesraad. Daaruit komt vooral het thema van nieuwe ontwikkelingen en de kansen of bedreigingen die deze ontwikkelingen opleveren naar voren. Wat bevoegdheden betreft komt uit de gesprekken duidelijk naar voren dat het daarbij vooral belangrijk is dat er een doorlopende dialoog is met de overheid, om te zorgen dat advies en beleid bij elkaar blijven komen. Aan de nadere vormgeving van de adviesraad, inbegrepen haar bevoegdheden, wordt nog gewerkt. Hierbij wordt er ook juridisch bekeken welke mogelijkheden de Kaderwet Adviescolleges hiertoe biedt.

De leden vinden de visie op het gebied van 'artificial general intelligence' nog weinig concreet. Kunt u nader toelichten wat Nederland doet op dit gebied? Ontwikkelt u hiervoor ook een strategie? (4)

³ De officiële term voor deze techniek is Retrieval Augmented Generation (RAG).

⁴ In de Engelse wettekst wordt gesproken van *General Purpose AI-modellen* (GPAI-modellen).

In de overheidsbrede visie op generatieve AI wordt onder 'artificial general intelligence' (AGI) het volgende verstaan: 'een technologie die intelligentie heeft over een breed scala aan domeinen waaronder redeneren, plannen en leren, en met deze vaardigheden op of boven het menselijke niveau presteert'. Er bestaat wetenschappelijke consensus over het feit dat AGI op dit moment nog niet is bereikt, maar de opvattingen over of en wanneer AGI precies zal worden bereikt lopen uiteen. Wel is het van belang om de komende jaren nauwgezet te volgen hoe de ontwikkelingen op het gebied van (generatieve) AI zich verder ontwikkelen. Daarom zijn in de actielijn 'nauwgezet volgen van alle ontwikkelingen' verschillende acties aangekondigd om op technologisch, socio-economisch en maatschappelijk niveau de ontwikkelingen te blijven monitoren.

Zoals ook aangegeven in het antwoord op voorgaande vragen van de fractie van GroenLinks-PvdA, moeten aanbieders van AI-modellen voor algemene doeleinden met systeemrisico's volgens de AI-verordening een modevaluatie uitvoeren en mogelijke systeemrisico's beperken. Onderdeel van de modevaluatie is ook het adversarial testen van het model. Dit is een evaluatiemethode die in een veilige omgeving de ingebouwde veiligheidsmechanismen van een AI-model uittest en probeert te omzeilen. Op deze manier kan inzichtelijk worden gemaakt hoe veilig een model is in de echte wereld.

Mochten er op een gegeven moment indicaties zijn dat er AI-modellen zijn die tekenen van 'artificial general intelligence' vertonen, dan spreekt het voor zich dat de mogelijke risico's hiervan als systeemrisico's kunnen worden gezien. Dit verplicht de aanbieder van het AI-model om deze risico's in kaart te brengen en te mitigeren.

Bovendien zien de leden twee sporen voor zich waarop Nederland een factor van belang kan zijn: het stimuleren en bijdragen aan Nederlandse of Europese techbedrijven die concurreren met de grote techgiganten om de machtsconcentratie op te breken, of het specialiseren in een gerichte toepassing van AI waar Nederland een niche mee invult en toonaangevend in wordt. Bent u het eens met deze analyse? (5)

Het kabinet deelt de analyse en heeft zich in nationaal en Europees verband ingezet om het AI-ecosysteem te verstevigen (zie antwoord hierna).

Welke inzet levert u op deze mogelijke twee sporen? Welk uniek specialisme heeft Nederland te bieden in het ontwikkelen van waardengedreven AI en de onderliggende infrastructuur? Hoe versterkt u deze expertise? Welke samenwerkingen met andere EU-lidstaten liggen hierin voor de hand? (6)

De inzet is gericht op het versterken van het waardengedreven AI-ecosysteem, onder meer via de publiek-private Nederlandse AI Coalitie (NLAIC). Deze coalitie is erop gericht om publiek-private samenwerking in het land te versterken en krachten te bundelen op het gebied van kennis, innovatie en vaardigheden. Dit moet Nederland internationaal positioneren als één AI-ecosysteem, met als doel het aanjagen van het verzilveren van kansen voor mensgerichte AI. Zeven regionale hubs zorgen voor het verbinden van innovatieve midden- en kleinbedrijven (mkb), kennisinstellingen en andere organisaties bij de Nederlandse AI-benadering. In afstemming met de NLAIC is het meerjarige AiNed Nationaal Groeifonds investeringsprogramma gericht op de versnelling van de ontwikkeling en toepassing van verantwoorde en mensgerichte AI, kennis, innovatie en talent voor AI in Nederland en kijken naar internationale/Europese samenwerking nodig is in de komende jaren te stimuleren. Momenteel werkt AiNed aan de oprichting van verschillende AI-innovatielabs, waarbij kansen voor generatieve AI worden geïncorporeerd. Dit zal waarschijnlijk AI-toepassingen omvatten voor de gezondheidszorg, energie, productie en mobiliteit. Voor de financiering van digitale innovatie op AI worden open calls van de Nederlandse Organisatie voor Wetenschappelijk Onderzoeken (NWO) en Rijksdienst voor Ondernemend Nederland (RVO) ingezet. Begin 2023 is ook het tienjarige AI-programma ROBUST gestart voor de realisatie van 17 nieuwe AI onderzoeksclubs en 170 promovendi op het gebied van betrouwbare machine learning, ter versterking van de expertise op het gebied van AI in Nederland. Om AI in het onderwijs te implementeren, is de ontwikkeling van een Nationaal Onderwijs Lab AI (NOLAI) gestart met €80 miljoen uit het Nationaal Groeifonds. Daarnaast hebben we naast deze initiatieven ook verschillende nationale financiële generieke instrumenten - de Wet Bevordering Speur- en Ontwikkelingswerk (WBSO) en de Mkb innovatiestimulering topsectoren (MIT) - waarbij we bedrijven, voornamelijk mkb-bedrijven, ondersteunen bij het gebruik van AI.

Europees lopen er verschillende initiatieven ter bevordering van de bouwstenen van AI in de EU: data, rekenkracht, algoritmen en talent.⁵ Samenwerking met andere EU-lidstaten vindt al plaats op het gebied van deze bouwstenen. Op bijvoorbeeld het gebied van high performance computing vindt de samenwerking plaats binnen het European High Performance Computing Joint Undertaking (EuroHPC JU), waar Nederland onderdeel van is. Een ander voorbeeld is het stimuleren van Europese taalmodellen, is Nederland onlangs lid geworden van de Alliantie voor Taaltechnologieën - Europees Digitaal Infrastructuur Consortium (ALT-EDIC).

Ook op het gebied van onderwijs zet het kabinet via programma's als Npuls en NOLAI in op het stimuleren van waardengedreven AI en ook in Europees verband bepleit het kabinet intensivering van de uitwisseling van expertise en Europese samenwerking gericht op de ontwikkeling van een Europees innovatie-ecosysteem gericht op AI in het onderwijs, met respect voor Europese publieke waarden.⁶

De leden van de GroenLinks-PvdA-fractie hebben opmerkingen en vragen over hoofdstuk 3. De leden lezen daarin een uitgebreide uiteenzetting van de kansen en risico's van generatieve AI, maar missen op meerdere vlakken een duidelijke keuze tussen die belangen. De staatssecretaris benoemt in haar brief dat de wenselijkheid van generatieve AI afhangt van de ontwikkeling, technologie en de intenties van de gebruiker. De leden vinden dat deze visie juist moet dienen om de bescherming tegen onwenselijk gebruik vast te leggen. Hoe zorgt u ervoor dat de risico's gemitigeerd worden en de kansen op een verantwoordelijke wijze worden benut? (7)

De AI-verordening vereist dat AI-systemen en -modellen veilig zijn en fundamentele rechten respecteren voordat ze op de markt komen of worden gebruikt in de EU. Burgers en consumenten kunnen er dan op vertrouwen dat wanneer belangrijke beslissingen met behulp van een AI-systeem worden genomen, dat dat systeem betrouwbaar is. Bedrijven die hoog-risico AI of AI-modellen voor algemene doeleinden kopen en inzetten kunnen straks erop vertrouwen dat deze AI betrouwbaar is en van een hoge kwaliteit. Hierdoor wordt de drempel om deze systemen in te zetten lager. Voor bedrijven die hoog-risico AI ontwikkelen wordt duidelijk aan welke eisen hun AI-systemen moeten voldoen (rechtszekerheid) en het stelt ze in staat om straks makkelijker hun AI-producten in de hele EU op de markt te brengen. Ook kunnen ze hun 'CE-markering' gebruiken voor hun wereldwijde marketing als kwaliteitsstempel. Als overheid zorgen wij er daarbij via verschillende wegen voor dat binnen de (generatieve) AI op een verantwoorde wijze kan worden ingezet. Bijvoorbeeld door middel van de ontwikkeling en uitrol van het Impact Assessment Mensenrechten en Algoritmes (IAMA), dat is ontwikkeld om tijdig vast te stellen of wet- en regelgeving is nageleefd en of (AI-)systemen in lijn zijn met publieke waarden. Instrumenten als de AI-verordening en de IAMA worden daarnaast opgenomen in het algoritmekader. Dit is een ondersteund instrument voor overheden waarin de belangrijkste normen en regels staan opgenomen die relevant zijn bij de (verantwoorde) inzet van AI en algoritmes.

Over de rol van generatieve AI als productietool willen de leden van de GroenLinks-PvdA-fractie weten hoe duidelijk wordt gemaakt wie de eigenaar van gegenereerde content is. Is dit de eigenaar van het AI-model, de personen wiens werk in de trainingsdata is gebruikt, of de eindgebruiker? Welke vorm van eigenaarschap vindt u wenselijk en hoe zorgt u ervoor dat dit is geborgd? (8)

Alleen voortbrengselen van de menselijke geest (originele werken die door een mens zijn gecreëerd) kunnen voor auteursrechtelijke bescherming in aanmerking komen. Het auteursrecht stelt makers in staat mee te delen in de opbrengst van de exploitatie van hun werken van letterkunde, wetenschap en kunst en biedt bescherming tegen aantasting van hun werken en hun reputatie. Naast dit rechtsvaardigheidsargument spelen overigens ook overwegingen van maatschappelijke wenselijkheid een rol. Het auteursrecht biedt makers namelijk een belangrijke stimulans meer werken te scheppen. Voortbrengselen van AI zijn geen voortbrengselen van de

⁵ Zie voor het standpunt van het kabinet ten opzichte van deze initiatieven het BNC-fiche 'Mededeling stimuleren van startups en innovatie in betrouwbare AI' (Kamerstuk 22112, nr. 3908) en 'Fiche: Verordening supercomputerinitiatief kunstmatige intelligentie' (Kamerstuk 22112, nr. 3907).

⁶ Kamerstuk 2023/2024, 21 501-34 nr. 418.

menselijke geest en komen op zichzelf beschouwd dan ook niet beschermd. Ik acht dit juist, omdat de voornoemde overwegingen van rechtvaardigheid en maatschappelijk nut geen rol spelen.

Als het gaat om het gebruik van werken van letterkunde, wetenschap of kunst voor de training van (generatieve) AI-modellen is het auteursrecht onverkort van toepassing. Het trainen van generatieve AI met werken van letterkunde, wetenschap of kunst kan waarschijnlijk onder de uitzondering op het auteursrecht inzake tekst- en datamining ex artikel 15o van de Auteurswet ressorteren. Die uitzondering regelt dat het maken van een reproductie van een werk in het kader van tekst- en datamining onder twee cumulatieve voorwaarden geen inbreuk maakt op het auteursrecht en in dat geval dan zonder voorafgaande toestemming van de maker of zijn rechtverkrijgenden is toegestaan. In de eerste plaats moet degene die de tekst- en datamining verricht rechtmatig toegang hebben tot het werk. In de tweede plaats moet het maken van zo'n reproductie ten behoeve van tekst- en datamining niet door de maker of zijn rechtverkrijgenden uitdrukkelijk op passende wijze zijn voorbehouden zoals door middel van machinaal leesbare middelen bij een online ter beschikking gesteld werk. Deze uitzondering vloeit rechtstreeks voort uit artikel 4 van de richtlijn auteursrecht in de digitale eengemaakte markt uit 2019.

Op dit moment is er daarbij nog geen voor deze kwestie relevante zaak aanhangig gemaakt bij het Hof van Justitie van de Europese Unie, waardoor deze rechter hier nog niet over heeft kunnen oordelen.

Deze leden zijn van mening dat de risico's dusdanig groot en maatschappelijk ontwrichtend zijn dat enige terughoudendheid op zijn plaats is. Zij pleiten eerst voor een verstrekkende regulering op de ongebreidelde marktwerking achter AI-modellen en stevige waarborgen dat deze het publieke belang dienen, voordat de kansen volop worden benut. De leden lezen over pilots met generatieve AI binnen de overheid op het gebied van dienstverlening. Kunt u toelichten welke generatieve AI-toepassingen nu worden gebruikt, getest of ontwikkeld bij de verschillende ministeries? Op welke wijze wordt de Kamer geïnformeerd over de reikwijdte en bevindingen van deze toepassingen? (9)

Via de Overheidsbrede Community of Practice Generatieve AI worden ervaringen gedeeld ten aanzien van onder andere de (verantwoorde) inzet van pilots binnen de overheid. Zo vinden er op dit moment uiteenlopende pilots plaats bij verschillende bestuurslagen. In de voortgangsrapportage, die eind dit jaar met uw Kamer zal worden gedeeld, wordt nader stilgestaan bij verschillende van deze pilots.

Bovendien biedt het Algoritmeregister een Rijksbreed inzicht. Door middel van regels en eisen in het algoritmekader wordt menselijke tussenkomst bij de inzet van AI-geassisteerde besluitvorming geborgd. Met alle departementen is afgesproken dat het algoritmeregister up-to-date is in 2025. Dit betekent dat dan ten minste alle hoog risico algoritmen geregistreerd zijn in het algoritmeregister. Hierover brengen departementen periodiek een planningsbrief uit. De eerstvolgende planningsbrief zal in het eerste kwartaal van 2025 worden verstuurd aan uw Kamer.

De leden benadrukken dat menselijke tussenkomst een randvoorwaarde is bij AI-geassisteerde besluitvorming die betrekking heeft op mensen. Hoe is dit geborgd? (10)

In de AVG is bepaald dat mensen in beginsel het recht hebben niet te worden onderworpen aan een uitsluitend op geautomatiseerde verwerking gebaseerd besluit, waaraan rechtsgevolgen zijn verbonden of dat hen anderszins in aanmerkelijke mate treft (artikel 22 AVG). De kerngedachte achter het verbod is dat niemand mag worden onderworpen aan de gevolgen van een besluit enkel en alleen op basis van kenmerken van een bepaalde groep waartoe diegene behoort. Op dit verbod bestaat een aantal uitzonderingen. Zo is geautomatiseerde individuele besluitvorming bijvoorbeeld wel toegestaan indien dat is bepaald in een wet, die ook voorziet in passende maatregelen ter bescherming van de rechten en vrijheden van de betrokkene. Artikel 40 van de Uitvoeringswet Algemene verordening gegevensbescherming (UAVG) bepaalt dat het verbod niet geldt, wanneer de geautomatiseerde individuele besluitvorming noodzakelijk is vanwege een wettelijke verplichting die op de verwerkingsverantwoordelijke rust of dit noodzakelijk is voor de vervulling van een taak van algemeen belang. De besluitvorming mag dan niet op basis van profilering plaatsvinden. Deze bepaling is bedoeld om geautomatiseerde afhandeling van massale, simpele gebonden beschikkingen toe te staan, waarbij menselijke tussenkomst van weinig meerwaarde zou zijn. Dat is

het geval bij de uitoefening van gebonden bevoegdheden, waarbij er zeer geringe beoordelingsruimte is, zoals het opleggen van een boete bij snelheidsovertredingen.

Verder gelden voor de (AI-geassisteerde) besluitvorming door bestuursorganen de bestuursrechtelijke waarborgen zoals neergelegd in de Algemene wet bestuursrecht, zoals het zorgvuldigheidsbeginsel, evenredigheidsbeginsel en motiveringsbeginsel. Ook dient de betrokkene in geval van geautomatiseerde besluitvorming op grond van artikel 13 AVG nuttige informatie over de onderliggende logica, het belang en de verwachte gevolgen van de verwerking van persoonsgegevens te krijgen.

De AI-verordening versterkt deze bestaande mechanismen door te eisen dat hoog-risico AI-systemen op een manier worden ontwikkeld die passend menselijk toezicht op deze systemen mogelijk maakt. De mate van menselijk toezicht is afhankelijk van de risico's van het systeem. Ook is het verplicht voor partijen die het AI-systeem inzetten (de gebruiksverantwoordelijken) om ook daadwerkelijk toezicht te houden op de inzet van het hoog-risico AI-systeem (artikel 26, lid 2) en om de mensen over wie een beslissing wordt gemaakt met het systeem hierover te informeren (artikel 26, lid 11).

De risico's zijn volgens de leden van de GroenLinks-PvdA-fractie verstrekkend en zwaarwegend. Zo is het mogelijk grootschalig schenden van mensenrechten door ondoordachte AI-modellen een grote zorg. De leden lezen dat er 'mogelijke' schendingen zijn van rechten op privacy, gegevensbescherming en auteursrechten. Echter is dit toch al een bewezen probleem van generatieve AI-modellen getraind op data via 'scraping'? (11)

Aan de vraag van de leden van de GroenLinks-PvdA-fractie lijkt de premisse ten grondslag te liggen dat het trainen van generatieve AI met van het Internet gescrepte data überhaupt niet zou zijn toegestaan. Die premisse is wat het auteursrecht betreft onjuist. De Auteurswet bevat een uitzondering voor tekst- en datamining (artikel 15o), zoals bepaald in artikel 4 van Richtlijn (EU) 2019/790 inzake auteursrechten in de digitale eengemaakte markt. Onder 'tekst- en datamining' wordt verstaan: een geautomatiseerde analysetechniek die gericht is op de ontleding van tekst en gegevens in digitale vorm om informatie te genereren, zoals, maar niet uitsluitend, patronen, trends en onderlinge verbanden. Naar de letter kan daaronder ook het trainen van generatieve AI vallen. Het maken van een reproductie van een werk van letterkunde, wetenschap of kunst is toegestaan voor tekst- en datamining. Dit geldt mits er rechtmatige toegang tot het werk is en de makers of hun rechtverkrijgenden dit niet uitdrukkelijk hebben verboden. Dit verbod kan bijvoorbeeld via machinaal leesbare middelen bij een online beschikbaar werk worden aangegeven. Scrapen is toegestaan, wanneer de voornoemde voorwaarden worden gerespecteerd. Anders levert scrapen een inbreuk op het auteursrecht op. Dat is echter niet zonder meer het geval.

Bent u het met de leden eens dat het essentieel is om auteursrechten van onder andere journalisten, kunstenaars en wetenschappers beter te beschermen en het ongebreideld scrapen van data dit nu onmogelijk maakt? Zijn zij nog de eigenaar van de AI-producties die hun werk gebruiken? (12)

Ik ben het eens met de leden van de GroenLinks-PvdA-fractie dat het belangrijk is het auteursrecht van makers van werken van letterkunde, wetenschap of kunst te respecteren. Ongebreideld scrapen kan dit onder druk zetten. Daarom is het van groot belang dat de AI-verordening is aangenomen. Deze verordening verplicht aanbieders om een intern beleid op te stellen voor de naleving van auteursrechten en een gedetailleerde samenvatting van de gebruikte trainingscontent openbaar te maken. Dit helpt rechthebbenden beter te controleren of de voorwaarden voor tekst- en datamining worden nageleefd. Rechthebbenden hebben zich hard gemaakt voor deze transparantieplichting en zijn tevreden met de invoering ervan.

Zoals eerder beschreven, kan het trainen van generatieve AI met werken van letterkunde, wetenschap of kunst soms toegestaan zijn op basis van de uitzondering op het auteursrecht voor tekst- en datamining. Zelfs dan mag de output van generatieve AI niet te veel lijken op het werk waarmee de AI is getraind, om inbreuk op het auteursrecht van de makers te voorkomen. Indien nodig, kunnen rechthebbenden juridische stappen ondernemen.

Ook is het van groot belang dat AI-modellen zonder bias opereren, maar dit is moeilijk te controleren wegens het 'black box'-effect van generatieve AI. Hoe dwingt u AI-ontwikkelaars tot transparantie? Welke rol speelt het Rijks AI-validatieteam daarin? Hoe ver strekt haar bevoegdheid? (13)

Het AI-validatieteam heeft geen directe bevoegdheid om AI-ontwikkelaars tot transparantie te dwingen. Dit betreft een team binnen de Rijksoverheid dat vooral innoveert, beleidsmakers adviseert en werkt aan technische oplossingen zoals tooling en benchmarks voor algoritmes en AI. In Nederland heeft de Autoriteit Persoonsgegevens (AP) deze bevoegdheid conform de AVG wel, indien het algoritmes of AI-systemen betreft die persoonsgegevens verwerken.

Daarnaast biedt de AI-verordening afdwingbare eisen voor transparante AI-modellen, waar Nederland zich tijdens de onderhandelingen voor heeft ingezet. Ontwikkelaars van AI-modellen voor algemene doeleinden zijn verplicht om technische documentatie over het model op te stellen en beschikbaar te houden voor toezichthouders en om informatie op te stellen voor ontwikkelaars die het AI-model in hun eigen AI-systeem willen integreren. Deze informatie moet inzicht geven in de capaciteiten en de beperkingen van het AI-model, waaronder bias.

Deze eisen kunnen worden afgedwongen door het AI-bureau ('AI Office'), die toezicht zal houden op de eisen aan AI-modellen voor algemene doeleinden. Er wordt daarnaast ook gewerkt aan praktijkcodes en geharmoniseerde normen die praktische invulling geven aan de transparantie-eisen.

De marktmacht van enkele monopolistische megabedrijven zien de leden van deze fractie als een van de meest fundamentele mankementen van het digitale domein, met verstreckende en vaak onwenselijke invloed op de manier waarop AI-producten werken en gebruikt worden. Deze leden vragen u hoe u dit weegt in deze visie: is het vergroten van de afhankelijkheid van buitenlandse techgiganten niet rechtstreeks in strijd met de breed gedragen wens voor digitale strategische autonomie binnen Nederland en Europa? (14)

Op 17 oktober 2023 is uw Kamer geïnformeerd over de Agenda Digitale Open Strategische Autonomie (DOSA)⁷. Deze agenda bouwt voort op het bredere kabinetsbeleid rond open strategische autonomie⁸ en aanpak strategische afhankelijkheden⁹.

Het kabinet hecht grote waarde aan een internationale rechtsorde waarin handel en investeringen vrijelijk hun weg vinden. Wederzijdse afhankelijkheden vormen de ruggengraat van het open handelssysteem en de internationale samenwerking waar Nederland en de EU veel profijt van hebben en die ons een sterkere geopolitieke positie opleveren. Tegelijkertijd kan een vorm van overheidsingrijpen nodig zijn, bijvoorbeeld indien strategische afhankelijkheden hoge risico's met zich meebrengen. Een integrale belangenafweging is vereist.

In de agenda DOSA staat de probleemanalyse en beleidsinzet op de beleidsprioriteit AI geschetst. Hier staat aangegeven dat grotere technologiebedrijven buiten de EU een veel betere uitgangspositie hebben voor de ontwikkeling van AI dan Europese bedrijven. Het gevolg hiervan is dat vele (overheids)organisaties en mkb afhankelijk kunnen worden van een kleine groep ontwikkelaars van (generatieve) AI.

Tegelijkertijd kan het stoppen van het afnemen van buitenlandse modellen en applicaties ertoe leiden dat Nederlandse en Europese organisaties niet de vruchten zullen plukken van de nieuwste technologie en dat innovatie die voortbouwt op de zogeheten foundation models niet hier zal plaatsvinden. Met als gevolg een verslechterde (digitale) concurrentiepositie en juist verminderde open strategische autonomie.

Het is belangrijk om te kijken naar in hoeverre het gebruik van buitenlandse technologie risico's oplevert en in hoeverre deze risico's (kunnen) worden gemitigeerd door huidig instrumentarium of aanvullend instrumentarium. Ook is het belangrijk om te kijken naar het versterken van de uitgangspositie van Europese AI-bedrijven. Hierop lopen al diverse acties, zoals staat geschetst in agenda DOSA en de overheidsvisie generatieve AI. Zo loopt in het kader van de EU Economische

⁷ Kamerstuk 2023/2024, 36 259, nr. 21.

⁸ Kamerstuk 2022/2023, 35 982, nr. 9.

⁹ Kamerstuk 2022/2023, 30 821, nr. 181.

Veiligheidsstrategie een AI-risicoanalyse, neemt Nederland deel aan het Europese partnerschap voor het verder ontwikkelen van supercomputing van wereldklasse (EuroHPC), en wordt het AI-ecosysteem via het Nationaal Groeifonds AiNed-programma versterkt.

De leden zijn tevens bezorgd over de invloed van AI op arbeid. Waarop baseert u de claim dat economen geen afname van werkgelegenheid verwachten? Is de huidige arbeidsmarkt voldoende voorbereid op de mogelijke verschuivingen van werk? In welke sectoren en welke soorten functies worden (op korte termijn) de grootste verschuivingen verwacht en hoe bereidt u werknemers daarop voor? (15)

De voorspelling dat AI op lange termijn geen grote afname op de totale werkgelegenheid zal veroorzaken is gebaseerd op empirisch onderzoek naar het arbeidsmarkteffect van eerdere technologische disrupties en op arbeidseconomische theorie. Tot nu toe is werkgelegenheid telkens toegenomen als gevolg van technologische ontwikkeling en automatisering.¹⁰ AI zal taken en daarmee (delen van) banen vervangen; maar de verwachting is dat ook nieuwe taken bijkomen - doordat nieuwe technologie nieuwe taken creëert en/of doordat stijgende welvaart de vraag naar arbeid doet groeien.

Veranderende werkinhoud als gevolg van AI zal wel om aanpassing door werkenden vragen. AI kan naar verwachting taken overnemen die draaien om het verwerken en analyseren van data. Dit beïnvloedt bijvoorbeeld het werk van professionals die medische, financiële of juridische gegevens analyseren. Het vermogen van generatieve AI om tekst, beeld en geluid te produceren, kan ook in de creatieve industrie voor verschuiving van werkinhoud zorgen. Werkenden zullen daardoor te maken krijgen met aanpassingen in hun werk. En in sommige gevallen zullen werkenden ook nieuw werk moeten vinden, en wellicht een periode werkloos zijn. De overheid biedt dan via de WW een vangnet. Daarnaast wordt door het kabinet in samenwerking met sociale partners, UWV en gemeenten de arbeidsmarktinfrastructuur versterkt, zodat mensen sneller en beter van werk naar werk kunnen worden begeleid. Daarnaast ligt er een opgave voor werkenden en werkgevers om te zorgen dat werkenden de juiste nieuwe vaardigheden opdoen.

Om meer zicht te krijgen op de impact van AI op arbeid en de sociaaleconomische implicaties heeft het kabinet de SER om advies gevraagd. Daarin komt ook de vraag terug welke kennis, vaardigheden en competenties werkenden nodig hebben om aansluiting op de arbeidsmarkt te behouden en hoe ze daarbij adequaat ondersteund kunnen worden. Dit advies wordt begin 2025 verwacht.

Op pagina 25 stelt u de volgende vragen: “Wie is er bijvoorbeeld verantwoordelijk voor schadelijke dan wel illegale inzet of output van bepaalde generatieve-AI-model? De modelontwikkelaar, de ontwikkelaar van de toepassing, de organisatie die de tool implementeert of de gebruiker?” Dit vinden de leden terecht vragen, maar zij verwachten dat het kabinet hier ook een uitgesproken mening over heeft. Kunt u de voor- en nadelen noemen van ieder van deze actoren? Welke rolverdeling vindt u wenselijk en hoe stuurt u hierop aan? (17)

Het kabinet vindt het belangrijk dat de verantwoordelijkheden worden belegd bij de partij die in staat is om deze op zich te nemen. In de AI-verordening worden verschillende verplichtingen opgelegd aan de actoren in de generatieve AI-keten.

Met betrekking tot schadelijke output, is het van belang dat aanbieders van AI-modellen voor algemene doeleinden met mogelijke systeemrisico's deze risico's vaststellen en beperken. Als dit niet voldoende gebeurt, bijvoorbeeld doordat het model nog steeds op grote schaal schadelijke content produceert, dan kan het AI-bureau maatregelen nemen om de aanbieder te dwingen om aan de eisen te voldoen of om het model van de markt te halen.

Daarnaast moet een aanbieder van een AI-model voor algemene doeleinden alle benodigde informatie geven aan de partij die dat model gebruikt in een generatieve AI-systeem, om te kunnen voldoen aan de AI-verordening. De aanbieder van het AI-model is immers de enige die beschikking heeft over deze informatie. Vervolgens moet de aanbieder van het generatieve AI-systeem ervoor zorgen dat de output van dit systeem wordt gelabeld als ‘kunstmatig gegenereerd’. Deze labels kunnen er in sommige gevallen voor zorgen dat de output van een model of systeem minder

¹⁰ Bijvoorbeeld: [Why Are There Still So Many Jobs? The History and Future of Workplace Automation on JSTOR](#), [The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment \(aeaweb.org\)](#)

schadelijk is, bijvoorbeeld in het geval van desinformatie.

De leden van de GroenLinks-PvdA-fractie hebben nog enkele opmerkingen over hoofdstuk 4. Dit hoofdstuk geeft een duidelijk overzicht van huidige wet- en regelgeving. Deze leden zijn echter benieuwd hoe de controle op deze wetgeving is geregeld. Zij zien dat de grote stroom aan nieuwe regelgeving zorgt voor tegenstrijdigheden en onduidelijkheid bij sommige besluiten. Welke toezichthouder is verantwoordelijk voor het beslechten van dergelijke interpretatieverschillen en het scheppen van precedent? Is er een toezichthouder die de samenhang van wetten bewaakt en een duidelijker handelingsperspectief ontwikkelt als houvast voor uitvoerders en bedrijven? (18)

Digitalisering en AI houden zich niet aan de grenzen van domeinen en sectoren. Dat betekent in de praktijk dat toezichtsvraagstukken steeds vaker bij verschillende toezichthouders terecht komen. Zo kan bijvoorbeeld een algoritme in de zorg waarin persoonsgegevens worden verwerkt het terrein zijn van zowel de Autoriteit Persoonsgegevens (AP) als de Inspectie Gezondheidszorg en Jeugd (IGJ). Toezichthouders kunnen dan onderling, bijvoorbeeld via een samenwerkingsconvenant, bepalen hoe er in een specifieke casus toezicht wordt gehouden en/of gehandhaafd. Daarbij vraagt het domeinoverstijgende karakter van AI om coördinatie en uitwisseling tussen (veel verschillende) toezichthouders, juist ook om de consistentie in het toezicht te bewaken en het stapelen van regels te voorkomen. Zo kunnen ondertoezichtgestelden veel baat hebben bij door toezichthouders gezamenlijk afgestemde normuitleg en guidance, zodat zij (vooraf) helderheid hebben over de gestelde kaders.

Coördinatie binnen het (digitale) toezichtsdomein is een doel uit de Werkagenda Waardengedreven Digitaliseren. Een belangrijk onderdeel hiervan is het versterken van de samenwerking tussen digitale toezichthouders. In Nederland zijn er verschillende toezichthouders die coördinerende taken op zich nemen. Zo hebben bijvoorbeeld de Directie Coördinatie Algoritmes (DCA) bij de AP en de Rijksinspectie Digitale Infrastructuur (RDI) verschillende soorten coördinerende taken op AI- en algoritmegebied. De DCA zet zich bijvoorbeeld in voor domeinoverstijgende risico-signalering op algoritmegebied en het komen tot concrete guidance voor ondertoezichtgestelden, zodat het vooraf helder is hoe en aan welke kaders zij moeten voldoen bij de inzet van AI en algoritmes. Ook richt de Algoritme- en AI-Kamer van het Samenwerkingsplatform Digitale Toezichthouders (SDT) zich op concrete afstemming over toezichtsvraagstukken op algoritme- en AI-gebied, waaronder mogelijk ook het uniform uitleggen van bepaalde normen. In deze Algoritme- en AI-Kamer zijn tal van toezichthouders binnen het digitale domein vertegenwoordigd.

De leden steunen de richting van 'open source, tenzij'. Het is hen alleen nog niet duidelijk wanneer de 'tenzij' van pas komt. Wanneer acht u het onmogelijk om open-source software in te zetten? Welke kaders geeft u organisaties die deze afweging moeten maken? (19)

Het 'tenzij'-voorbehoud heeft betrekking op situaties of gevallen waarin open-source (te veel) afdoet aan de veiligheid van het betreffende AI-systeem en/of misbruik faciliteert. Deze risico's komen voort uit het feit dat er bij open-sourcemedellen geen centrale controle is op aanpassingen van het oorspronkelijke model. Kwaadwillenden kunnen hier hun voordeel mee doen door ingebouwde veiligheidsmechanismen te omzeilen.

De afweging tussen 'open source' en het voorbehoud 'tenzij' rust op een aantal factoren. Dit is ten eerste de robuustheid van de veiligheidsmechanismen in het model die het genereren van illegale content en/of het gebruiken van auteursrechtelijk beschermd materiaal verhinderen. Ten tweede is de toepassingscontext van belang, m.b.t. misbruik- en cybersecurityrisico's en de mogelijke impact daarvan.

Daarnaast wordt er gewerkt aan herziening van het Rijksbreed Cloudbeleid.¹¹ Hierbij wordt onder meer aandacht besteed aan de strategische afhankelijkheden van clouddiensten en de doorgifte van persoonsgegevens en overheidsdata naar landen buiten de EU.

De leden bedanken u en de betrokken ambtenaren voor het beantwoorden van de vragen.

Vragen en opmerkingen van de leden van de VVD-fractie en antwoord

¹¹ Kamerstuk 2023/2024, 26 643, nr. 1149.

De leden van de VVD-fractie hebben kennisgenomen van de Overheidsbrede visie op generatieve AI (AI). Deze leden hebben nog enkele vragen en opmerkingen.

De leden van de VVD-fractie lezen dat generatieve AI kan bijdragen aan de duurzaamheidstransitie. Hoe zullen op korte termijn de zorgen over het energiegebruik opgevangen worden? Ook met oog op de krapte op het stroomnet? (1)

Recent is de aandacht voor AI sterk toegenomen. Naar aanleiding hiervan leven er allerlei zorgen over de impact van AI op energieverbruik. Dit is een zeer terechte zorg. De ambities van Nederland om binnen de EU koploper te zijn op het gebied van verantwoorde (generatieve) AI-toepassingen, vallen of staan ook bij voldoende oog voor belangrijke aspecten als energie(verbruik) en duurzaamheid. Dat betekent dat we als overheid deze technologie niet inzetten als deze grote schade aanricht aan mens en planeet. Daarbij moet wél worden meegenomen dat netwerk- en computerapparatuur zich tot nu toe altijd razendsnel heeft mee-ontwikkeld. Denk bijvoorbeeld aan chips en processoren. Onder andere met de inzet van slimme toepassingen op basis van AI kan het energiegebruik van dergelijke apparatuur, maar bijvoorbeeld ook van energienetten, verder geoptimaliseerd worden. Daarnaast is het verband tussen datagroei en energieverbruik voor veel onderdelen van de digitale infrastructuur vaak kleiner dan verwacht. Zo leidt een stijging van dataverkeer met één procent niet tot een even grote toename van de energieconsumptie. De krapte op het stroomnet zal daarbij op de korte termijn een gegeven zijn waar we mee moeten omgaan. Hoe AI zal uitpakken op de energievraag in Nederland en wereldwijd, laat zich op dit moment niet goed voorspellen. Het ministerie van Economische Zaken en Klimaat houdt goed de vinger aan de pols. Ook laat dit ministerie dit jaar een onderzoek uitvoeren naar de duurzame impact van nieuwe AI-modellen. Zo kunnen we op nationaal, Europees en internationaal niveau bijsturen als dat nodig is.

De bovengenoemde leden vinden het goed dat de opgave om mensen te ondersteunen om zich te blijven ontwikkelen om vaardigheden te vergaren die aansluiten op de arbeidsmarkt mocht hun baan verdwijnen door technologie wordt aangepakt. Is hier ook een dialoog met het ministerie van SZW over? (2)

Deze dialoog vindt zeker plaats. De overheidsbrede visie is ook mede namens de Minister van Sociale Zaken en Werkgelegenheid verstuurd. Ook hebben BZK, EZK en SZW nauw samengewerkt aan de adviesaanvraag 'AI, de toekomst van werk en sociaaleconomische implicaties' aan de SER, die gelijktijdig met de overheidsbrede visie generatieve AI aan uw Kamer is verzonden.¹²

De leden van de VVD-fractie delen de zorgen over de afhankelijkheid van taalmodellen uit niet-Europese landen. Kunnen deze modellen naast het stroken met onze normen en waarden ook bijdragen aan discriminatie? Hoe wordt dit voorkomen? (3)

Taalmodellen die in niet-Europese landen worden ontwikkeld maar wel in de Europese Unie op de markt worden gebracht, moeten aan de eisen uit de AI-verordening voldoen. Als het, volgens de AI-verordening, om AI-modellen voor algemene doeleinden met systeemrisico's gaat, zal de ontwikkelaar van een dergelijk model ervoor moeten zorgen dat ook risico's zoals grootschalige discriminatie worden voorkomen.

Wordt er ook ingezet op de ontwikkeling van taalmodellen die in Europa ontwikkeld worden? (4)

Om de ontwikkeling van Nederlandse taalmodellen te stimuleren is recent 13,5 mln. euro vanuit EZK/RVO toegekend voor de ontwikkeling van GPT-NL, een NL-taalmodel gebaseerd op publieke waarden en voor publieke doelen. Hiermee wordt in 2024 een open en transparant Nederlands taalmodel ontwikkeld. In 2025 en latere jaren kunnen eerste toepassingen worden ontwikkeld met gebruik van rekenkracht van de nationale supercomputer (Snellius) gehost door SURF (organisatie verantwoordelijk voor de ICT-infrastructuur voor onderzoek en onderwijs).

Voor het stimuleren van Europese taalmodellen, is Nederland sinds kort lid van de Alliance for Language Technologies - European Digital Infrastructure Consortium (ALT-EDIC). AiNed is gemandateerde partij, TNO ondersteunt. Het doel van de ALT-EDIC is om een gemeenschappelijke

¹² [SER-adviesaanvraag 'AI, de toekomst van werk en sociaaleconomische implicaties' | Rapport | Rijksoverheid.nl](#)

Europese infrastructuur in taaltechnologieën op te richten. Zo kunnen we doorbouwen op de resultaten van het nationale GPT-NL via Europese samenwerking.

Ook is het van belang om de ontwikkeling van taalmodellen voor talen zoals Fries en Papiamentu te stimuleren. Vanuit BZK wordt daarom een bijdrage verkend aan een Fries taalmodel dat bijvoorbeeld kan worden ingezet voor de verbetering van de overheidsdienstverlening.

Deze leden lezen dat de Autoriteit Persoonsgegevens toezicht moet houden op de naleving van de Algemene Verordening Gegevensbescherming (AVG). Is er genoeg ruimte en capaciteit binnen deze organisatie om dit goed uit te voeren? (5)

De Autoriteit Persoonsgegevens (AP) krijgt door het coalitieakkoord structureel extra gelden om het toezicht op de AVG te versterken. Het budget van de AP is in 2024 een lumpsum bedrag van ongeveer 41 miljoen euro. Het is van belang dat de AP haar taken op een goede manier kan (blijven) uitvoeren. Deze vraag wordt dan ook meegenomen in de gesprekken met deze toezichthouder.

De leden van de VVD-fractie vragen op welke manier een AI-adviesorgaan (of 'Rapid Response Team') meegenomen kan worden in de besluiten die het kabinet neemt over belangrijke ontwikkelingen op het gebied van (generatieve) AI. (6)

Momenteel verkennen de ministeries van BZK, EZK en OCW de mogelijkheid van een dergelijk adviesorgaan rondom AI, dat de overheid zou kunnen adviseren met betrekking tot de ontwikkelingen op het vlak van AI. Meerdere experts benadrukken in de gesprekken die hierover gevoerd worden het belang van de doorlopende dialoog met de overheid, om te zorgen dat advies en beleid bij elkaar aansluiten. Wat de advisering van het kabinet betreft, geldt dat wanneer een dergelijk adviescollege een rol krijgt bij het adviseren over wetgeving of beleid, de regels van de Kaderwet Adviescolleges van toepassing zijn. Afhankelijk van de gewenste rol van een dergelijk adviesorgaan ten opzichte van het kabinet zal moeten worden bekeken wat hierin gewenst is.

Tot slot willen de leden van de VVD-fractie graag de aangenomen motie Rajkowski en Dekker-Abdulaziz (Kamerstuk 30821-193) in herinnering brengen. Eerder heeft het kabinet in een brief aangegeven het verbieden van ontwrichtende aanbevelingsalgoritmes niet uit te gaan voeren. Tegelijkertijd lezen deze leden dat het kabinet momenteel de inrichting van een AI-adviesorgaan aan het verkennen is. Wanneer dit AI-adviesorgaan operationeel is, zouden de leden van de VVD-fractie graag de aangenomen motie Rajkowski en Dekker-Abdulaziz (Kamerstuk 30821-193) voor nader advies willen voorleggen aan het te vormen AI-adviesorgaan. Is het kabinet, gelet op de brede meerderheid voor de motie, voornemens de motie voor te leggen aan het in te richten AI-adviesorgaan? Zo nee, waarom niet? (7)

Zoals hierboven aangegeven is het onderzoek naar de mogelijkheden van een AI-adviesorgaan nog in een verkennende fase. Dat geldt ook voor de inrichting en de vraag aan wie dit adviesorgaan adviseert: de overheid in brede zin of het kabinet zelf. Uit de gesprekken die hierover gevoerd worden komt duidelijk naar voren dat er een sterke aansluiting moet zijn tussen de adviezen van dit orgaan en het beleid dat gemaakt wordt rondom AI. Indien dit orgaan een adviesfunctie richting de overheid krijgt waarbij beleidsmakers in gesprek kunnen met experts rondom AI-thema's, is het mogelijk om daaraan een vraag voor te leggen over aanbevelingsalgoritmes, zoals benoemd in de aangenomen motie¹³.

Vragen en opmerkingen van de leden van de NSC-fractie en antwoord

De leden van de NSC-fractie hebben met interesse kennisgenomen van de overheidsbrede visie op generatieve AI.

De leden van de NSC-fractie hechten eraan het volgende te benadrukken. Het is met het oog op de ontwikkelingen van de afgelopen jaren niet onwaarschijnlijk dat AI als systeemtechnologie wereldwijd voor nog onvoorzienbare veranderingen gaat zorgen op mondiaal niveau. Dit is afhankelijk van de wijze en het tempo waarin generatieve en overige vormen van AI ontwikkelen, evenals de toepassingsgebieden waarvoor de technologie inzetbaar blijkt en de systemen waarop AI opereert. Een cruciale kwestie is of (en op welke termijn) uiteindelijk een vorm van 'general intelligence' met een hoog adaptief en evoluerend vermogen zal kunnen ontstaan. Dit zou in

¹³ Kamerstuk 2022/2023, 30 821, nr. 193.

negatieve scenario's tot gevolg kunnen hebben dat de mens de grip op de technologie verliest, met gevolgen in het spectrum van maatschappelijk verstorend tot existentieel bedreigend. Hierbij is onder meer te denken aan – al dan niet intentioneel – sabotage van (vitale) infrastructuur, digitale massasurveillance en een betrouwbaarheids crisis in het domein van informatievoorziening. Er bestaat geen consensus over de waarschijnlijkheid van deze scenario's, maar een significant aantal toonaangevende AI-experts heeft reeds aangegeven de potentiële dreiging serieus te nemen en heeft in dat kader opgeroepen tot actie en internationale coördinatie. Dit was onder andere de aanleiding voor de eerste AI Safety Summit van afgelopen november. De overheidsbrede visie benoemt ook het belang van een actieve rol voor Nederland op het internationale toneel, met onder andere het bepleiten van 'pre-deployment audits' en het beschikbaar stellen van de 'model weights' van de grote, geavanceerde modellen. De leden van deze fractie zouden graag zien dat die visie op waardengedreven en specifiek veilige AI in een internationale context verder uitgediept wordt (1).

De inzet van de regering is erop gericht om de juiste balans aan te brengen ten aanzien van zowel de kansen als de risico's die AI biedt. De voornaamste inzet van de regering richt zich daarom op de implementatie van de AI-verordening, waarvan de verwachting is dat die een grote bijdrage zal leveren aan het mitigeren van risico's. Nederland neemt daarnaast deel aan de AI Safety Summit (AISS) overleggen over AI veiligheid. Op de eerste in die reeks committeerden een aantal van de grote AI-bedrijven zich eraan om toegang te verlenen tot hun volgende generaties modellen voor zgn. 'pre-deployment audits.' Ondanks dat ik niet aan de tweede bijeenkomst in Seoel kon deelnemen ben ik verheugd te constateren dat enkele van die bedrijven inderdaad zulke toegang verleend hebben aan verschillende AI Safety-instituten. De regering kijkt uit naar de volgende editie, de AI Action Summit, die voor februari 2025 in Parijs gepland staat. Daar zal naar verwachting ook het International AI Safety Report worden gepresenteerd, waarvan onlangs de interimversie op de AISS-top in Seoel werd gelanceerd.

Naast de invoering van de AI-verordening en de opvolging van AISS komt het onderwerp veilige AI ook aan de orde in verschillende processen bij de Verenigde Naties (VN), zoals het werk van de High-level Advisory Body on AI (HLAB-AI). Het kabinet steunt het werk van deze zeer diverse adviesraad, die uitvoerig heeft geconsulteerd over diens tussentijdse aanbevelingen. De laatste van die consultaties heb ik georganiseerd, op het thema AI en democratie, mensenrechten en de rechtstaat. De regering kijkt uit naar het eindrapport dat in augustus verwacht wordt, en waarin concrete aanbevelingen verwacht worden ten aanzien van te nemen stappen op mondiaal niveau om onder meer de veiligheid van AI te borgen. Parallel hieraan wordt momenteel het Global Digital Compact onderhandeld, waarin eveneens voorstellen gedaan worden op het gebied van AI governance. Hierbij richt de Nederlandse inzet zich wederom op het vinden van de juiste balans tussen kansen en risico's. Overigens wordt ook gewerkt aan de uitvoering van de motie Six Dijkstra¹⁴, over het in kaart brengen van de rechtsstaat bedreigende AI-ontwikkelingen en weerbaarheidsbevorderende maatregelen.

Met de overweging dat enerzijds de doorontwikkeling van hoogwaardige AI mogelijk risico's van een ongekende aard en omvang met zich meebrengt, en dat anderzijds de belangrijkste ontwikkelingen op AI zich hoofdzakelijk buiten de landsgrenzen afspelen, is het de opvatting van de leden van de NSC-fractie dat Nederland moet komen tot een risicogebaseerde internationale AI-strategie, vergelijkbaar met de Internationale Cyberstrategie (ICS). Met het besef dat AI-incidenten niet categorisch voorkomen zullen kunnen worden, achten de leden het daarnaast van belang dat adaptieve maatregelen voor de Nederlandse fysieke en digitale leefomgeving in kaart gebracht worden met als doel potentiële schade door toekomstige geavanceerde AI te mitigeren of voorkomen.

Het kabinet stelt, in de ogen van de leden van de NSC-fractie, terecht dat de keten van de ontwikkeling van state-of-the-art AI-modellen een grote afhankelijkheid kent van Amerikaanse marktpartijen, maar dat Nederland vanwege zijn sterke halfgeleiderecosysteem tegelijkertijd een unieke positie in deze keten heeft. Dit maakt dat de leden benieuwd zijn wat de inzet van het kabinet gaat zijn binnen de Europese samenwerking om een goede concurrentiepositie op AI voor niet alleen de Europese Unie, maar ook specifiek voor Nederland te realiseren. Zij verzoeken het kabinet dit nader toe te lichten. (2)

¹⁴ Kamerstuk 2023/2024, 26 643, nr. 1151. |

Het kabinet heeft zich ervoor ingezet om zowel nationaal als in Europees verband de concurrentiepositie op AI te versterken. De grote investeringen die gemaakt moeten worden om bijvoorbeeld de hoeveelheid data en rekenkracht te verkrijgen, genoodzaakt een EU-aanpak. Daarom heeft het kabinet ook het onlangs gelanceerde AI-innovatiepakket van de Europese Commissie verwelkomd, waar onder andere wordt voorgesteld om de rekencapaciteit en data in Europa toegankelijker te maken voor startups. Ook wordt er gewerkt aan de andere bouwstenen die nodig zijn voor een sterke concurrentiepositie op AI, zoals talent en vaardigheden. Nederland draag actief bij aan verschillende lopende Europese initiatieven ter bevordering van de concurrentiepositie op AI, onder andere door nationale cofinanciering voor projecten op het gebied van dataruimten voor energie en de maakindustrie, de ontwikkeling van een Europees open-source foundation model en de ALT-EDIC.

Daarnaast wordt actief gewerkt aan het versterken van het nationale AI-ecosysteem en daarmee de Nederlandse concurrentiepositie op AI, onder meer via het Nationaal Groeifonds AiNed-programma en de publiek-private Nederlandse AI Coalitie (NLAIC). Voor de financiering van digitale innovatie op AI worden ook open calls van de Nederlandse Organisatie voor Wetenschappelijk Onderzoeken (NWO) en Rijksdienst voor Ondernemend Nederland (RVO) ingezet. Om ook de drempel tot rekencapaciteit voor geavanceerde AI-modellen te verkleinen voor bedrijven, met in het bijzonder mkb, wordt ook de mogelijkheid voor een AI-faciliteit met high performance computing verkend.

De leden van de NSC-fractie vragen, nu het kabinet benoemt dat significante investeringen in AI-infrastructuur benodigd zijn, of het kabinet kan inschatten of de huidige Nederlandse datacenter- en netcapaciteit volstaat om een volwassen AI-ecosysteem te creëren en te onderhouden, en zo niet, welke keuzes en afwegingen hierin gemaakt moeten worden. (3)

De beantwoording van deze vraag komt terug in de antwoorden op vraag 14 van de GroenLinks-PvdA-fractie (gedeelte over het AI-ecosysteem) en vraag 1 van de VVD-fractie (gedeelte over netcapaciteit).

Het kabinet stelt dat de hogere arbeidsproductiviteit ten gevolge van generatieve AI met het oog op historische trends zou kunnen leiden tot een toename in materiële welvaart, betere gezondheid en meer vrije tijd. De leden van de NSC-fractie willen hier wat kritische kanttekeningen bijplaatsen, omdat dit in hun ogen neigt naar een papieren werkelijkheid. Deze leden stellen ten eerste dat productiviteitsgroei niet de facto ook leidt tot een brede en eerlijke verdeling van de welvaart. Ten tweede willen zij benoemen dat het historisch gezien zeker niet evident is dat verhoogde productiviteit maakt dat mensen, specifiek diegenen die werken in dienstverband, meer vrije tijd krijgen. De leden verzoeken daarom het kabinet deze voorspelling nader te onderbouwen. Omdat het visiedocument hier niet nader op ingaat, vragen de leden welke maatregelen het kabinet wil treffen om te stimuleren dat productiviteitsgroei door AI leidt tot brede welvaart en een betere kwaliteit van leven. (4)

De overheidsbrede visie stelt dat generatieve AI kansen biedt voor het verhogen van brede welvaart en de kwaliteit van leven. Bij een gelijkblijvende productiviteit is er – simpel gesteld – sprake van een afruil tussen zaken als besteedbare inkomens, publieke voorzieningen en vrije tijd omdat meer van het één ten koste gaat van het ander. Bij een groeiende productiviteit hoeft een toename van het één niet ten koste van het ander te gaan. Productiviteitsgroei is niet hetzelfde als brede welvaart, maar biedt daar zo bezien wel meer mogelijkheden voor een toename hiervan.

Het kabinet herkent het punt dat productiviteitsgroei niet vanzelfsprekend leidt tot een verbetering van brede welvaart. En dat hierbij ook de vraag speelt hoe de baten van productiviteitsgroei worden verdeeld. Onderhandelingen tussen werkgevers- en werknemersorganisaties bepalen in belangrijke mate het aandeel van de productiviteitsgroei dat naar werkenden gaat. Daarnaast vindt er herverdeling van welvaart plaats via het stelsel van fiscaliteit en toeslagen en de uitkeringen. Het gesprek over de mate en manier van herverdeling vindt continu plaats, binnen het kabinet en met uw Kamer. Ook bij de verdeling van de baten van toekomstige productiviteitsgroei zullen politieke keuzes aan de orde zijn over de gewenste mate van herverdeling.

De studie van het CPB naar de Nederlandse economie in historisch perspectief toont een trend op lange termijn richting een steeds kortere werkweek, parallel aan toenemende welvaart en

productiviteit.¹⁵ Deze studie toont dus een toename in vrije tijd. Met het oog op de krappe arbeidsmarkt en vergrijzing ziet het kabinet overigens ook de noodzaak voor werkenden om waar mogelijk meer uren te werken.

Het kabinet heeft de SER om advies gevraagd over de impact van AI op de arbeidsmarkt en de sociaaleconomische implicaties. Hierin wordt ook aandacht besteed aan de relatie tussen AI en productiviteit. Het advies van de SER wordt begin 2025 verwacht.

De leden van de NSC-fractie vragen of er inmiddels iets meer bekend is over de contouren van het eventuele Rapid Response Team AI. Deze leden vragen het kabinet wat het maakt dat dit initiatief nog steeds in de verkenningsfase zit en welke afwegingen hierin nog spelen. (5)

Graag verwijs ik voor de beantwoording van deze vraag naar de antwoorden op de vragen van de VVD-fractie over de mogelijkheden van een AI-adviesorgaan.

Vragen en opmerkingen van de leden van de D66-fractie en antwoord

De leden van de D66-fractie hebben met interesse kennisgenomen van de overheidsbrede visie op generatieve AI. Deze leden hebben nog enkele vragen naar aanleiding van het visiestuk en de begeleidende brief.

Allereerst moedigen de leden van de D66-fractie van harte aan dat er ook goed wordt gekeken naar de kansen en mogelijkheden die generatieve AI biedt aan de samenleving. Gelukkig wordt dat ook veelvuldig benoemd in zowel de brief als het visiestuk. Zo wordt AI bijvoorbeeld gezien als 'probleemoplosser' voor problemen binnen de gezondheidszorg. Deze leden zien graag een specificering op deze terreinen. Op welke manier kan AI als probleemoplosser functioneren voor maatschappelijke thematiek (zoals de gezondheidszorg) en wat er is er voor nodig om dat te bewerkstelligen? (1)

De beschikbare (wetenschappelijke) literatuur wijst uit dat er reeds geëxperimenteerd wordt met generatieve AI en diverse maatschappelijke opgaven. Zo worden er voor het onderwijs verschillende (complexere) taken benoemd die generatieve AI kan vervullen, zowel voor leerlingen en studenten als voor docenten. Voorbeelden zijn samenvattingen maken, lesmateriaal schrijven, lessen inplannen en het op maat maken van lesstof. Verder kan generatieve AI helpen bij het beoordelen van het werk van leerlingen en studenten. Zo zijn veelgenoemde (potentiële) kansen voor het gebruik van generatieve AI in het onderwijs tijdwinst en efficiënte, hogere kwaliteit van de lesstof, betere leeropbrengst en het bevorderen van de motivatie van leerlingen.¹⁶

Er bestaan ook kansen voor generatieve AI in de gezondheidszorg. Zo kan generatieve AI voor patiënten als hulpmiddel fungeren, door bijvoorbeeld ondersteuning te bieden bij de uitleg van complexe diagnostiek in de vorm van virtuele gezondheidsassistenten of het vertalen naar begrijpelijk taalniveau. Daarnaast kan generatieve AI ook worden ingezet voor de ondersteuning bij het uitvoeren van administratieve taken, bijvoorbeeld door het opstellen van patiëntrapporten gebaseerd op patiëntdata en het samenvatten van gespreksverslagen.¹⁷ Om hier goed zicht te krijgen heeft de Minister van Volksgezondheid, Welzijn en Sport (VWS) een eerste inventarisatie op de kansen en risico's voor generatieve AI laten opstellen. Zie voor een nadere toelichting over AI binnen het zorgdomein ook het antwoord op de volgende vraag van de leden van D66.

Dat de leden van de D66-fractie zich afvragen wat ervoor nodig is om bovenstaande kansen te bewerkstelligen is dan ook een zeer terechte en belangrijke vraag. Of een bepaalde capaciteit van generatieve AI zich uiteindelijk als kans of als risico manifesteert, hangt af van de specifieke ontwikkeling, de toepassing van de technologie en de intenties dan wel expertise van de gebruiker.

Het is daarom belangrijk om adequaat regie te voeren en toezicht te houden op de kansen én risico's van generatieve AI. Zo wordt op basis van de overheidsbrede visie op generatieve AI onderzoek naar de inzet van generatieve AI met het oog op verschillende maatschappelijke vraagstukken gestimuleerd, zoals klimaatopgaven en democratie. Tegelijkertijd wordt ingezet op

¹⁵ <https://www.cpb.nl/de-nederlandse-economie-historisch-perspectief>

¹⁶ https://www.rathenau.nl/sites/default/files/2023-12/Scan_Generatieve_AI_Rathenau_Instituut.pdf

¹⁷ Meskó, B., & Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ digital medicine, 6(1), 120.

initiatieven die de ecologische voetafdruk van generatieve AI verminderen en het risico op desinformatie verlagen.

Het derde uitgangspunt, 'het dienen en borgen van menselijke autonomie', moedigt enerzijds de inzet van AI in de zorg aan, en geeft aan dat het tegelijkertijd onwenselijk kan zijn. Vervolgens is de conclusie dat er maatschappelijk debat nodig is. Deze leden zouden graag een duidelijker standpunt zien ten opzichte van het zorgveld, waarbij de inzet – mits volgens alle voorzorgmaatregelen - wordt aangemoedigd en aangejaagd, zeker gezien de opgave binnen de zorg. (2)

Het kabinet onderschrijft dat AI als systeemtechnologie een kansrijke digitale innovatie is die oplossingen kan bieden voor structurele vraagstukken in de zorgsector. Zoals eerder beschreven in de brief 'Waardevolle AI voor gezondheid'¹⁸, vraagt dit om een goede balans tussen grip op de zorgvuldige totstandkoming en toepassing van AI én de flexibiliteit om te kunnen experimenteren en leren. Hier is ook met uw Kamer over gesproken in de vorm van het schriftelijk overleg 'Informatie en Communicatietechnologie in de Zorg' op 25 augustus 2022.¹⁹ Om AI in de zorg te stimuleren zijn de afgelopen jaren de eerste stappen dus reeds genomen. Daarnaast worden de komende periode de huidige risico's en kansen in het zorgveld geïnventariseerd.

Vanuit het VWS-programma 'Waardevolle AI voor gezondheid' zijn tussen 2020 en 2022 reeds meerdere producten ontwikkeld die het zorgveld ondersteunen bij het ontwikkelen en implementeren van AI-systemen. Zo is er de Leidraad Kwaliteit AI in de Zorg²⁰, de Handleiding Aanpak begeleidingsethiek voor AI in de zorg²¹, het Hulpmiddel Handelingsruimte²² en het Hulpmiddel Mindmap Bestuurlijk agenderen²³. Deze producten bieden nog steeds handvatten voor de inzet van AI in het zorgveld. Om aan te blijven sluiten op de technologische en juridische ontwikkelingen rondom AI, zet het Kabinet ook de komende periode in op het ondersteunen van het zorgveld bij het implementeren en opschalen van passende AI. Zo wordt onderzocht welke behoefte er in de zorgsector bestaat aan ondersteuning rond de aankomende AI-verordening, waarin op basis van risico-classificatie regels worden gesteld die moeten gaan helpen bij het inzetten van passende AI. Ook is een eerste inventarisatie uitgevoerd naar de kansen en risico's voor generatieve AI in de zorg en is de Nederlandse AI Coalitie gevraagd om de voorbereidingen te treffen om learning communities op te zetten. Deze learning communities moeten de zorgsector ondersteunen in het opbouwen en uitwisselen van kennis op het gebied van AI. Daarnaast wordt naar verwachting voor het einde van het jaar het standpunt op ethiek van AI in de zorg met uw Kamer gedeeld door de Minister van Volksgezondheid, Welzijn en Sport.²⁴

De leden van de D66-fractie constateren dat er al veel initiatieven zijn gestart om de vertaalslag van generatieve AI te maken naar constructieve bijdragen aan de samenleving. Zo bestaan er Ethical Legal and Societal Apects (ELSA)-labs waarin onderzoek wordt gedaan naar ethische en juridische aspecten van AI en worden er innovatielabs georganiseerd in samenwerking met AiNed. Kan het kabinet een aantal concrete resultaten of winstpunten benoemen die deze initiatieven hebben opgeleverd, behalve de samenwerkingsverbanden zelf? (3)

Begin 2023 zijn de eerste vijf ELSA Labs van start gegaan op het gebied van maatschappelijke vraagstukken, waar meer dan 70 organisaties aan deelnemen. Wat doet AI bij de productie en verspreiding van desinformatie, zoals deepfakes, en de rol van de journalistiek, is één van de vragen waar het ELSA Lab AI, Media & Democracy zich in verdiept. Dit is onder andere relevant voor het ontwikkelen en gebruik van algoritmen die nieuwsaanbevelingen doen. Momenteel staat een nieuwe ELSA Lab call open met de vraag wat de ELSA implicaties zijn van huidige en toekomstige AI-relevante wetgeving, zoals de AI-verordening, voor specifieke toepassingsgebieden

¹⁸ Kamerstuk 27 529, nr. 275.

¹⁹ Kamerstuk 27 529, nr. 281.

²⁰ <https://guideline-ai-healthcare.com/>

²¹ <https://www.datavoorgezondheid.nl/documenten/publicaties/2021/01/15/handleiding-aanpak-begeleidingsethiek-voor-ai--digitale-zorg>

²² <https://www.datavoorgezondheid.nl/documenten/publicaties/2021/07/15/hulpmiddel-handelingsruimte-waardevolle-ai-voor-gezondheid>

²³ <https://www.datavoorgezondheid.nl/documenten/publicaties/2021/10/12/hulpmiddel-mindmap-bestuurlijk-agenderen-waardevolle-ai-voor-gezondheid>

²⁴ Kamerstuk 32805, nr. 176.

zoals gezondheid en zorg en de energietransitie. Via een netwerkproject wordt de komende jaren ook tussen de labs geleerd hoe mensgerichte AI oplossingen aan te pakken.

Momenteel werkt AiNed nog aan de oprichting van verschillende AI-innovatielabs, waarbij kansen voor generatieve AI worden geïncorporeerd. Dit zal waarschijnlijk AI-toepassingen omvatten voor de gezondheidszorg, energie, productie en mobiliteit.

De leden van de D66-fractie juichen toe dat de overheid investeert om verschillende sectoren voor te bereiden op het gebruik van generatieve AI. Uit de brief blijkt dat er voor een periode van tien jaar wordt gefinancierd met middelen vanuit het Nationaal Groeifonds. Echter, in hoeverre liggen deze middelen al vast? Kan deze financiering door een volgend kabinet weer worden stopgezet, of zijn deze verplichtingen juridisch verplicht? (4)

In de brief wordt ingegaan op verschillende initiatieven gefinancierd uit het Nationaal Groeifonds (NGF). Voor het meerjarige AiNed programma is € 160,5 mln uit het NGF toegekend. Daarnaast is er € 44 mln. voorwaardelijk toegekend. De Groeifondscommissie is gevraagd om voor de zomer een advies uit te brengen over de definitieve toekenning van deze € 44 mln. Het Nationaal Groeifondsprogramma Digitaliseringsimpuls onderwijs NL (Npuls) ontwikkelt onder meer een landelijk AI-punt en AI-visie voor het mbo, hbo en wo om zo de sectoren voor te bereiden op de transformatie van het onderwijs met AI en samen met partners en instellingen deze veranderingen ook mede vorm te geven. Voor dit project is € 560 miljoen toegekend uit het Nationaal Groeifonds in 2022. Daarvan is € 140 miljoen onvoorwaardelijk toegekend en € 420 miljoen voorwaardelijk toegekend. Voorwaardelijke toekenningen zijn net als de directe toekenningen juridisch verplichte middelen, die al door het parlement zijn geautoriseerd. Wanneer er wordt voldaan aan gestelde voorwaarden, kan het kabinet op basis van een advies van de adviescommissie de voorwaardelijke toekenning omzetten in een definitieve toekenning.

Er is uit het NGF € 79,6 miljoen toegekend voor het Nationaal Onderwijslab AI (NOLAI). In dit project worden prototypes ontwikkeld van digitale intelligente onderwijsinnovaties om de kwaliteit van het primair en voortgezet onderwijs te verbeteren. NOLAI onderzoekt in één van de co-creatieprojecten hoe het onderwijs regie houdt over het leren en onderwijzen bij het gebruik van generatieve AI en wat de consequenties zijn op pedagogisch-didactisch, ethisch, technologisch en professionaliseringsvlak. Er is naast deze toekenning een reservering van € 63 miljoen om deze prototypes door te ontwikkelen tot producten die geschikt zijn om breed in de markt en in de klas in te zetten en tegelijkertijd het hele ecosysteem te versterken. Ten aanzien van reserveringen heeft communicatie plaatsgevonden en zijn bestuurlijke verwachtingen gewekt, maar deze middelen zijn niet juridisch verplicht. Voor het omzetten van een reservering in een (voorwaardelijke) toekenning zal de adviescommissie op basis van een vernieuwd of aangevuld investeringsvoorstel een nieuw advies uitbrengen aan de fondsbeheerders. De omzetting van een reservering wordt door de fondsbeheerders aan de Ministerraad voorgelegd. De omzettingen van de reserveringen zijn definitief na autorisatie door beide Kamers.

Uit de brief blijkt ook dat het overheidsbreed de inzet is om nieuwe generaties generatieve AI met Nederlandse en Europese partners te ontwikkelen. In hoeverre zijn daar financiële middelen aan gekoppeld en zijn daar al voorbeelden van, zo vragen de leden van de D66-fractie? (5)

Er is recent 13,5 mln. euro vanuit EZK/RVO toegekend voor de ontwikkeling van GPT-NL, een Nederlands taalmodel voor publieke doelen. In 2024 wordt een open en transparant Nederlands taalmodel ontwikkeld. In 2025 en latere jaren kunnen eerste toepassingen worden ontwikkeld met gebruik van rekenkracht van de nationale supercomputer (Snellius) gehost door SURF (organisatie verantwoordelijk voor de ICT-infrastructuur voor onderzoek en onderwijs).

De leden van de D66-fractie lezen dat er duidelijk wordt gemaakt waar AI-toepassingen volgens deze visie aan moet voldoen en waar ontwikkelaars zich aan dienen te houden. Tegelijkertijd weten we dat generatieve AI veel van wat online staat kan 'scrapen' en daar trainingsmodellen mee kan vullen, ook als dit bijvoorbeeld persoonsgegevens bevat. In hoeverre zijn toezichthouders in staat om te deduceren met welke data taalmodellen zijn getraind, ook gezien de constatering in de overheidsvisie zelf dat generatieve AI modellen een significant 'black-box karakter hebben' en 'een gebrek aan interpreteerbaarheid hebben'? Welke methodes hebben toezichthouders om bij generatieve AI-modellen die *niet* open-source werken in te zien of de taalmodellen met de juiste data zijn getraind? (10)

In artikel 58 van de AVG zijn de bevoegdheden van de Autoriteit Persoonsgegevens (AP) opgenomen, waaronder de bevoegdheid om de verwerkingsverantwoordelijke en de verwerker op te dragen alle noodzakelijke informatie aan de AP te verstrekken (artikel 58, eerste lid, aanhef en onder a, AVG). Ook heeft de AP bij de verwerkingsverantwoordelijke en de verwerker toegang tot alle persoonsgegevens en noodzakelijke informatie, evenals tot alle bedrijfsruimten met inbegrip van alle uitrustingen en middelen voor gegevensverwerking (artikel 58, eerste lid, aanhef en onder e en f, AVG). Daarnaast zijn ook de in titel 5.2 van de Algemene wet bestuursrecht neergelegde bevoegdheden van toepassing, zoals het vorderen van inlichtingen (artikel 5:16 Awb), het vorderen van inzage in zakelijke gegevens en bescheiden (artikel 5:17 Awb) en het onderzoeken van zaken (artikel 5:18 Awb). De verwerkingsverantwoordelijke en de verwerker zijn verplicht alle medewerking te verlenen aan de AP (artikel 5:20 Awb). Op grond van deze bevoegdheden kan de AP onderzoek verrichten naar de data waarmee generatieve AI-modellen zijn getraind.

Daarnaast bevat de op 21 mei jl. door de Raad aangenomen AI-verordening regels voor AI-modellen voor algemene doeleinden, waaronder generatieve AI-modellen zijn begrepen (zie titel VIII A). Zodra deze wet in werking treedt, moeten deze modellen aan bepaalde transparantievereisten voldoen, zoals het publiceren van gedetailleerde samenvattingen van de trainingscontent waarmee de modellen zijn getraind.

De leden van de D66-fractie maken zich zorgen om het auteursrecht dat gepaard gaat met generatieve AI-modellen. In de overheidsbrede visie wordt daar op ingegaan. Het is nu nog niet duidelijk of generatieve AI-modellen een beroep kunnen doen op de tekst- en datamining-exceptie. Dit betekent dat er mogelijk al lange tijd auteursrechten worden geschonden. Deze leden willen graag weten of er geen stappen mogelijk zijn om daar meer zekerheid in te krijgen, zoals het inwinnen van advies van de landsadvocaat of de Raad van State? (11)

De leden van de aan het woord zijnde fractie merken terecht op dat op dit moment nog niet met volledige zekerheid valt te zeggen dat het trainen van generatieve AI kan profiteren van de uitzondering op het auteursrecht inzake tekst- en datamining. Dit vergt een authentieke interpretatie van artikel 4 van de richtlijn auteursrecht in de digitale eengemaakte markt waarop die uitzondering is gebaseerd en daartoe is uiteindelijk alleen het Hof van Justitie van de Europese Unie bevoegd. De interpretatie van de landsadvocaat en de Raad van State is daarvoor dus niet doorslaggevend. In de AI-verordening uit 2024 wordt er in ieder geval wel vanuit gegaan dat het trainen van AI de uitzondering inzake tekst en datamining valt. De transparantieverplichting uit de AI-verordening strekt er namelijk toe rechthebbenden beter in staat te stellen te controleren of de voorwaarden van de uitzondering inzake tekst- en datamining worden nageleefd.

De leden begrijpen dat online beschikbare werken nu alleen zijn beschermd als er een correct voorbehoud is gemaakt. In hoeverre is dit bekend, zo vragen de leden van deze fractie, en zou dit niet breder gecommuniceerd moeten worden? (12)

Zoals ook uiteengezet in de beantwoording van vraag 8 van de fractie van Groenlinks-PvdA, is het auteursrecht in een online omgeving onverkort van toepassing. Het trainen van generatieve AI met werken van letterkunde, wetenschap of kunst kan waarschijnlijk onder de uitzondering op het auteursrecht inzake tekst- en datamining ex artikel 15o van de Auteurswet ressorteren. Die uitzondering regelt dat het maken van een reproductie van een werk in het kader van tekst- en datamining onder twee cumulatieve voorwaarden geen inbreuk maakt op het auteursrecht en in dat geval dan zonder voorafgaande toestemming van de maker of zijn rechtverkrijgenden is toegestaan. De voorwaarden hiervoor worden in de beantwoording van vraag 8 van de Groenlinks-PvdA-fractie (nader) benoemd.

De verplichtingen uit de recent aangenomen AI-verordening, waarbij aanbieders van AI-modellen voor algemene doeleinden (zoals generatieve AI) onder meer een gedetailleerde samenvatting van de trainingscontent openbaar moeten maken en een intern beleid voor auteursrechten moeten opstellen, sluiten aan op het hierboven geschetste juridisch kader in relatie tot het auteursrecht (zie ook beantwoording vraag 12 GroenLinks-PvdA-fractie).

Er is daarbij onderzoek verricht naar en gepubliceerd over de uitzondering voor tekst- en datamining, met name ook over de wijze waarop een voorbehoud kan worden gemaakt.²⁵ In de praktijk zijn daarvoor allereerste nieuwe standaarden in ontwikkeling, bijvoorbeeld in het kader van

²⁵ Laatstelijk: A.P. Engelfriet & D.J.G. Visser, *Werk de mijnwerk optout voor mijn werk?*, *Auteursrecht*, 2024/1.

de European Digital Infrastructure Consortium, in aanvulling op bijvoorbeeld robot.txt waarop ten tijde van de onderhandelingen voorafgaande aan de totstandkoming van de eerdergenoemde richtlijn in 2019 vooral oog bestond. Het toepasselijk juridisch kader is bekend en in de praktijk worden nieuwe en verfijndere standaarden ontwikkeld om voorbehouden te kunnen maken.

Ook wordt er gesteld dat output door AI alleen auteursrechtelijk wordt beschermd als er een samenwerking tussen een mens en AI is, maar vervolgens wordt gesteld 'waarschijnlijk is de menselijke invloed op het uiteindelijke resultaat nog niet voldoende om van een werk in een de auteursrechtelijke betekenis van het woord te kunnen spreken'. Deze leden zouden dit graag verhelderd zien worden: betekent dit dat door AI gegenereerd werk sowieso nooit een 'werk in de auteursrechtelijke betekenis van het woord' kan zijn, ondanks de hoeveelheid creatieve keuzes? Wie maakt daarin de afweging? (13) En in hoeverre is dit onderdeel van de AI Act? (14)

Het auteursrecht geeft makers van werken van letterkunde, wetenschap of kunst het voor een rechter afdwingbare recht om anderen te verbieden of toe te staan daarvan gebruik te maken. Het Europese Hof van Justitie heeft in het zogenaamde Infopaq-arrest²⁶ beslist dat werken een schepping van de geest zijn. Het materiaal moet oorspronkelijk zijn in die zin dat het om een eigen intellectuele schepping van de maker ervan. Voorbeelden van werken zijn boeken, films en muziek. Bescherming bestaat tegen "verveelvoudigen" (zoals kopiëren, vertalen, bewerken) en "openbaar maken" (zoals uitvoeren, uitzenden, verspreiden van exemplaren, ter beschikking stellen op het Internet).

Zonder menselijke tussenkomst kunnen voortbrengselen van generatieve AI dus niet voor auteursrechtelijke bescherming in aanmerking komen. Er is namelijk geen sprake van een schepping van de menselijke geest. Als er sprake is van samenwerking tussen een mens en een artificieel intelligent algoritme kan het voortbrengsel onder omstandigheden voor auteursrechtelijke bescherming in aanmerking komen. Noodzakelijk daarvoor is dat de vorm het resultaat is van scheppende menselijke arbeid en dus van creatieve keuzes als voortbrengsel van de menselijke geest. De creatieve keuzes kunnen in de voorbereidende fase worden gemaakt. Daarbij kan worden gedacht aan de uitwerking van het idee in de prompt van de AI-applicatie. In de beleidsbrief wordt betwijfeld of de menselijke invloed op het uiteindelijke voortbrengsel daarmee op zichzelf beschouwd voldoende is om van een werk in de auteursrechtelijke betekenis van het woord te kunnen spreken. Dat kan veranderen door later creatieve keuzes toe te voegen bij de bewerking van het voortbrengsel van AI tot het uiteindelijk werk dat openbaar wordt gemaakt. In de menselijke bewerking van de output van generatieve AI kunnen voldoende creatieve keuzes zijn gemaakt die tot uitdrukking zijn gekomen in het uiteindelijke resultaat.

Bij de beantwoording van de vraag of sprake is van een prestatie die voor auteursrechtelijke bescherming in aanmerking komt, gaat het dus om de menselijke creativiteit, de gemaakte originele keuzes, het persoonlijke stempel van de maker, tot uiting komende in het uiteindelijke resultaat. De kwaliteit van dat resultaat is niet ter zake doende. Of er sprake is van een auteursrechtelijk beschermd werk zal, als van generatieve AI gebruik wordt gemaakt, afhangen van de omstandigheden van het geval en staat in geval van een geschil uiteindelijk ter beoordeling aan de rechter. De AI-verordening heeft daarmee niets van doen.

Tot slot merken de leden van de D66-fractie op dat in de zes actielijnen van de kabinetsbrede visie weinig staat over het aantrekkelijk maken en stimuleren van het ondernemingsklimaat voor start-up/scale-ups op het gebied van generatieve AI. In hoeverre ligt daar een duidelijke visie aan ten grondslag en zijn daar financiële middelen aan gekoppeld? In hoeverre is er een duidelijk Europees initiatief, waarbij wordt geïnvesteerd in veelbelovende innovaties (zoals bij het Groeifonds)? (15)

Voor startups en scale-ups zijn al tal van generieke en financiële instrumenten beschikbaar, zoals WBSO, Innovatiekrediet, InvestNL en MIT-regeling en daarnaast de private investeringsfondsen voor risicodragend kapitaal. Ook heeft TeachLeap.NL een katalysatorrol om startups naar de scale-up fase te brengen. Dit levert op het gebied van (generatieve) AI mooie resultaten op. Zo zijn er via InvestNL de afgelopen jaren bijvoorbeeld investeringen gedaan in AI-gedreven tech-ondernemingen

²⁶ <https://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=CELEX:62008CJ0005>.

met een impactvolle toepassing in sectoren zoals onderwijs, digitale gezondheidszorg en klimaattechnologie.²⁷

*Bedrijven die zich richten op (generatieve) AI kunnen ook van dit bestaande instrumentarium gebruik maken. Daarnaast lopen er programma's als het NGF AiNed om ook iets extra's te kunnen doen voor AI-gedreven mkb bedrijven, waaronder startups. Door de Europese Commissie zullen ook extra mogelijkheden worden geboden voor start- en scale-ups, zoals het beschikbaar stellen van rekenkracht voor grote AI-modellen.*²⁸

Vragen en opmerkingen van de leden van de BBB-fractie en antwoord

De leden van de BBB-fractie hebben de stukken ter kennis aangenomen en hebben hierover de volgende vragen;

De leden van de BBB-fractie wijzen op de volgende passage uit de brief (Kamerstuk 26643-1125): "Het kabinet richt de aandacht in het bijzonder op de impact van deze nieuwe digitale technologie op de samenleving." (blz. 1). Hierover vragen deze leden wat de vastgestelde grenzen zijn aan het gebruik van AI met betrekking tot de bevoegdheden van overheden die ingrijpen in de rechtspositie van mensen. (1)

Bij overheidsingrijpen in de rechtspositie van mensen met het gebruik van AI zijn allereerst dezelfde regels en vastgestelde grenzen van toepassing, als bij ander overheidsingrijpen in de rechtspositie van mensen. Zo gelden er voor besluitvorming algemene beginselen van behoorlijk bestuur, zoals het in de Algemene wet bestuursrecht gecodificeerde zorgvuldigheidsbeginsel, evenredigheidsbeginsel en motiveringsbeginsel. Bovendien is overheidsingrijpen alleen toegestaan als er geen of een gerechtvaardigde inbreuk is op grondrechten. Gebruik van AI door overheden mag bijvoorbeeld niet in strijd zijn met het recht op gelijke behandeling en het recht op eerbiediging van de persoonlijke levenssfeer. Ten aanzien van de verwerking van persoonsgegevens heeft de AVG regels gesteld die ook voor het gebruik van AI door overheidsinstanties gelden. Belangrijke beginselen voor verwerking van persoonsgegevens zijn onder andere transparantie, rechtmatigheid, behoorlijkheid en doelbinding. Verder bepaalt de AVG dat betrokkenen in beginsel het recht hebben om niet onderworpen te worden aan een uitsluitend geautomatiseerde besluitvorming (artikel 22 AVG). Ook gelden er specifieke regels voor de inzet van AI, zoals de recentelijk aangenomen AI-verordening die eisen en kaders bevat voor de ontwikkeling en het gebruik van AI-systemen door overheden.

De leden van de BBB-fractie lezen het volgende: "Ook bestaan er zorgen over mogelijke schendingen van rechten op het gebied van privacy, gegevensbescherming en auteursrechten." Kan de staatssecretaris aangeven in hoeverre de mogelijkheid tot bijvoorbeeld 'scrapen' een bedreiging vormt voor de onafhankelijke en kwalitatieve nieuwsvoorziening in Nederland? (2)

De leden van de aan het woord zijnde fractie vragen aandacht voor het standpunt van de brancheorganisatie voor nieuwsbedrijven NDP Nieuwsmedia die van mening is dat de impact van generatieve AI op de nieuwsvoorziening onderbelicht is gebleven in de kabinetsvisie. Het kabinet heeft wel degelijk oog voor mogelijke inbreuken op het auteursrecht door ontwikkelaars van generatieve AI. Het maken van een kopie van perspublicaties van uitgevers om generatieve AI te trainen is op grond van het geldende, op de richtlijn auteursrecht in de digitale eengemaakte markt gebaseerde recht alleen zonder voorafgaande toestemming van de persuitgever toegestaan op grond van de uitzondering inzake tekst- en datamining, wanneer de kopie van de perspublicatie wordt gemaakt ten behoeve van tekst- en datamining, er rechtmatige toegang tot de perspublicatie bestaat en de uitgevers geen uitdrukkelijk voorbehoud heeft gemaakt. Verder mogen de gemaakte kopieën niet worden bewaard, maar moeten zij na de tekst- en datamining meteen

²⁷ Zie bijvoorbeeld: <https://www.invest-nl.nl/actueel/invest-nl-investeert-in-capital>

²⁸ Zie ook de BNC-fiches 'Mededeling stimuleren van startups en innovatie in betrouwbare AI' (Kamerstuk 22112, nr. 3908) en 'Verordening supercomputerinitiatief kunstmatige intelligentie' (Kamerstuk 22112, nr. 3907).

worden verwijderd. Het geven van authentieke interpretatie aan die uitzondering is uiteindelijk voorbehouden aan het Hof van Justitie van de Europese Unie. De recent tot stand gekomen AI-verordening verplicht de aanbieders van generatieve AI om een voldoende gedetailleerde samenvatting van de gebruikte trainingscontent openbaar te maken. Dit stelt uitgevers in staat te controleren of er inderdaad rechtmatige toegang tot de perspublicaties bestaat en de geplaatste voorbehouden van de uitgevers worden gerespecteerd. NDP Nieuwsmedia vindt die verordening een belangrijke stap in de goede richting. Het ligt op de weg van de uitgevers van perspublicaties zich goed te organiseren teneinde het persuitgeversrecht daadwerkelijk en effectief uit te oefenen. In het buitenland worden inmiddels afspraken gemaakt over het gebruik van auteursrechtelijk beschermde werken om ai te trainen.²⁹ De handhaving van dat recht, in geval van inbreuk, is niet strafrechtelijk gesanctioneerd en dus een zuiver privaatrechtelijke aangelegenheid. De mogelijkheden die de in het nationale recht omgezette Europese handhavingsrichtlijn daarvoor biedt zijn daarvoor afdoende.

Deze leden lezen tevens: “Wat werk betreft, zijn er zorgen over de impact van generatieve AI op werkgelegenheid, de kwaliteit van werk en inkomensverdeling. Zowel op de korte als lange termijn kan generatieve AI leiden tot werkloosheid.” (blz. 5). Kan de staatssecretaris verduidelijken welk deel van de arbeidspopulatie mogelijk geraakt gaat worden door de invloed van AI, hoeveel procent dit is, welke sectoren het meest positief en negatief beïnvloed gaan worden door de opkomst van AI? Hoeveel mensen zouden er cijfermatig omgeschoold moeten worden omdat een baan geautomatiseerd wordt, en hoeveel is dit procentueel van de totale arbeidspopulatie? (3)

Het Internationaal Monetair Fonds (IMF) wijst erop dat ongeveer 60 procent van de banen in hoogontwikkelde economieën, zoals Nederland, blootgesteld wordt aan AI.³⁰ Ruim de helft van die blootgestelde banen kent een hoge complementariteit, wat wil zeggen dat ze taken omvatten die we naar verwachting niet willen of kunnen automatiseren. In banen met een lage complementariteit en een hoge blootstelling aan AI bestaat een hoger risico op een lagere vraag naar arbeid en lagere lonen. Werknemers in banen met een hoge complementariteit en hoge blootstelling aan AI kunnen juist het meest profiteren van het gebruik van AI in hun werk.

Het is nu niet zeker vast te stellen welke banen de komende jaren geautomatiseerd worden en om hoeveel mensen dit zou gaan. Naast de huidige blootstelling en complementariteit hangt dat af van hoe de technologie zich precies verder ontwikkelt, hoe bedrijven deze adopteren, en hoe zij bestaande werknemers daarin meenemen.

In samenwerking met sociale partners, UWV en gemeenten wordt de arbeidsmarktinfrastructuur verstevigd, zodat mensen sneller en beter van werk naar werk kunnen worden begeleid. Daarnaast ligt er een opgave om te zorgen dat werkenden de juiste nieuwe vaardigheden opdoen, waarbij er ook een verantwoordelijk is bij voor werkgevers en werkenden zelf.

In de adviesaanvraag van het kabinet aan de SER, over 'AI, de toekomst van werk en sociaaleconomische implicaties' wordt daarbij ook aandacht besteed aan het vraagstuk werkgelegenheid en kwaliteit van werk in relatie tot (generatieve) AI. Het SER-advies wordt begin 2025 verwacht.

Kan de staatssecretaris aangeven waar de kansen liggen voor de Nederlandse economie in brede zin? (4)

Meer algemeen is AI een systeemtechnologie die de Nederlandse economie als geheel productief kan transformeren. Ongeacht de sector biedt de technologie bijzondere kansen om bij te dragen aan de verhoging van de Nederlandse productiviteit en de verlichting van arbeidstekorten. Marktpartijen zijn in beginsel het best in staat om de precieze kansen van AI van sector tot sector te herkennen en verkennen. Dit geldt zowel voor ondernemers die nieuwe toepassingen ontwikkelen en aanbieden, als ook voor partijen die er gebruik van maken.

²⁹ Zie bijvoorbeeld: <https://informatieprofessional.nl/reddit-sluit-contentovereenkomst-om-ai-model-te-trainen>.

³⁰ <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>

Tegelijk is al duidelijk dat AI in sommige sectoren kansen biedt die beter zijn te verzilveren met enige hulp van uit de overheid. Zo wordt in de overheidsbrede visie op generatieve AI ingegaan op de kansen van generatieve AI binnen verschillende sectoren in Nederland, waaronder de maakindustrie en ICT-sector. Ook in de Nationale Technologiestrategie worden kansrijke toepassingsgebieden genoemd die relevant zijn voor de Nederlandse economie, zoals de chipindustrie, landbouw- en voedselproductie, en de energietransitie. In al deze sectoren kunnen AI en data ingezet worden om processen te optimaliseren en daarmee de productiviteitsgroei te vergroten.

Daarnaast achten we het belangrijk Nederland in bredere zin in goede positie te brengen om de kansen in een AI-economie te benutten. Ook daartoe worden in de overheidsbrede visie op generatieve AI actielijnen geformuleerd. Denk aan beleidsinzet op: (1) het nauwgezet volgen van relevante ontwikkelingen omdat veel van de economische verandering nog onzeker is; (2) de vergroting van kennis en kunde van mensen om de kansen van AI daadwerkelijk te kunnen verzilveren; en (3) het innoveren met generatieve AI, met daarbij ook een experimenterende overheid die de unieke kans op verhoging van productiviteit in het publieke domein helpt realiseren.

De leden van de BBB-fractie citeren de passage: “Maar er is ook een risico dat het taken doet verschromelen, de autonomie van werkenden inperkt en de menselijkheid van werkrelaties onder druk zet.” (blz. 5). Kan de staatssecretaris aangeven hoe de autonomie van werkenden ingeperkt worden, en welk beleid er wordt gemaakt om dit te voorkomen? Kan de staatssecretaris verduidelijken hoe de menselijkheid van werkrelaties onder druk worden gezet en hoe dit voorkomen kan worden, en of hier actief beleid wordt gemaakt? (5)

AI-automatisering en verdringing van taken die door mensen worden uitgevoerd kunnen het gevoel van autonomie en de sociale verbondenheid op het werk doen afnemen. Autonomie komt in het geding als werknemers geen of minder grip op hun werk ervaren, zoals ook geconstateerd door de WRR en het onlangs verschenen Rathenau/TNO-rapport over algoritmisch management. Ook in Europa is daar aandacht voor. In de Europese richtlijn platformwerk waarover op 11 maart 2024 een akkoord bereikt is, is opgenomen dat werkenden niet via een app ontslagen mogen worden zonder menselijke tussenkomst. Een ander risico is dat sociale verbondenheid afneemt als menselijke interactie wordt vervangen door interactie met technologie, die de menselijke werknemer vervangt.

In de adviesaanvraag van het kabinet aan de SER, over 'AI, de toekomst van werk en sociaaleconomische implicaties' wordt expliciet aandacht besteed aan de vraag hoe geborgd kan worden dat de inzet van (generatieve) AI geen afbreuk doet aan de kwaliteit van werk, en deze waar mogelijk zelfs vergroot. Autonomie en sociale verbondenheid zijn daarbij aangemerkt als essentiële onderdelen van kwaliteit van werk. Het SER-advies wordt begin 2025 verwacht.

Deze leden lezen: “Hier ligt onder andere een opgave van digitale inclusie.” (blz. 5). Kan de staatssecretaris dit verder uitwerken, wat wordt er bedoeld met ‘digitale inclusie’, en hoe omschrijft de staatssecretaris dit? (6)

Digitale inclusie is het streven dat iedereen kan meekomen in de digitale samenleving. Zonder ingrijpen kan digitalisering mensen namelijk verder op achterstand zetten. Dat geldt met name voor laaggeletterden, mensen met een lager inkomen en senioren. Recent onderzoek laat zien dat ook AI digitale ongelijkheid kan vergroten. Daarom wordt ingezet op beleid voor digitale inclusie, zoals het bieden van ondersteuning, het vergroten van toegang tot digitale apparaten en internet en het bevorderen van digitale vaardigheden en digitaal bewustzijn. Ook wordt ingezet op het vergroten van digitale toegankelijkheid van websites en apps, zodat deze gebruikt kunnen worden door mensen met een beperking. Met Alliantie Digitaal Samenleven werk ik aan deze opgave samen met publieke, maatschappelijke en private partners.

De leden van de BBB-fractie zien de volgende passage: “Als alle mensen handvaten geboden wordt om nieuwe technologie te begrijpen en gebruiken, kunnen zij ook beter inspelen op de veranderingen die dat met zich meebrengt” (blz. 5). Hoe kunnen burgers die niet zo digitaal vaardig

zijn meekomen in de omgang met deze technologie? Wordt er behalve aan kinderen ook aan ouderen gedacht wat betreft het laten begrijpen en gebruiken van deze nieuwe technologie? (7)

Om burgers te leren over de werking van AI en de gevolgen hiervan lopen verschillende initiatieven vanuit de Werkagenda Waardengedreven Digitaliseren. Zo zijn 6 maart jl. de conceptkerndoelen voor de curriculumherziening van primair, voortgezet en (voortgezet) speciaal onderwijs gepubliceerd.³¹ Eén van de conceptkerndoelen is het verkennen van de werking van AI en de mogelijkheden en beperkingen van deze technologie. Nadat de kerndoelen in de schoolpraktijk zijn beproefd, zal de procedure starten om ze vast te leggen in wet- en regelgeving. Wanneer de kerndoelen zijn vastgesteld, krijgen alle leerlingen hier onderwijs in.

Via het programma Digitaal Burgerschap krijgen bibliotheken een rol om mensen actief, vaardig en weerbaar te maken in de online informatiesamenleving. Bijvoorbeeld met workshops en lezingen over de kansen en gevaren van AI. Het programma ondersteunt bibliotheken die hiermee willen starten of experimenteren. In het kader hiervan vond 23 januari jl. de landelijke AI-dag voor bibliotheken plaats, waar onder meer de AI-Inspiratiegids werd gepresenteerd.³²

Een ander initiatief om een algemeen publiek te informeren over de kansen en gevaren van AI is de Nationale AI Cursus.³³ Dit is een laagdrempelige online cursus waarin AI op een begrijpelijke manier wordt uitgelegd. Deze cursus is het resultaat van een privaat initiatief dat medegefinancierd is door het ministerie van BZK. Er is vanuit de Nationale AI Cursus ook een variant ontwikkeld voor kinderen en voor specifieke toepassingsgebieden: Zorg, AgriFood, Energie, Onderwijs, de Creatieve Industrie en Logistiek & Maritiem.³⁴ De Nationale AI Cursus wordt momenteel doorontwikkeld met modules over generatieve AI.

Specifiek voor ouderen ondersteunt het ministerie van BZK het programma Welkom Online van het Nationaal Ouderenfonds. Dit programma ontwikkelt en faciliteert workshops om mensen mee te nemen in de digitale samenleving, waarbij de behoefte van de deelnemer centraal staat. De inzet van nieuwe technologieën en het creëren van bewustzijn daaromtrent is hiervan een onderdeel. Ook de app Digihandig richt zich op ouderen. BZK ondersteunt deze app via het Startup in Residence-programma. Via deze app kunnen mensen oefenen om hun smartphone te gebruiken. Dit jaar wordt de app doorontwikkeld voor kritische digitale vaardigheden en nieuwe technologieën, zoals AI.

Deze leden lezen: “Een andere uitdaging betreft machtsconcentratie. De opkomst van generatieve AI-technologie versterkt de al bestaande dynamiek in digitale markten, waarmee het risico op misbruik van marktmacht toeneemt.” (blz. 5). Kan de staatssecretaris weergeven of en hoe mkb-bedrijven kunnen meekomen met het gebruik van deze technologie, en waar deze bedrijven nu tegenaan lopen bij het gebruik van deze nieuwe technologie? (8)

Er is inmiddels al een groot aantal startups bezig om specifieke toepassingen te ontwikkelen op basis van de bestaande modellen voor generatieve AI op basis van learning by doing. Omdat AI relatief nieuw is en mkb vaak de diepe AI-kennis (nog) niet in huis hebben, zijn er instrumenten als MIT-regeling om mkb-bedrijven te helpen bij het samenwerken met andere partijen op het gebied van onderzoek en innovatie, zoals mogelijk gemaakt via het NGF programma AiNed. Uiteraard moet dat gebeuren binnen de spelregels van de AI-verordening voor verantwoorde AI. Om bedrijven te ondersteunen in complexe compliance-vraagstukken met betrekking tot de AI-verordening, worden zogenaamde regulatory sandboxes ontwikkeld. Organisaties kunnen in deze regulatory sandboxes met toezichthouders in gesprek gaan om tot antwoorden op deze vraagstukken over regelgeving te komen. Deze antwoorden worden vervolgens weer breder gedeeld, zodat ook andere organisaties deze kennis kunnen gebruiken.

³¹ <https://www.rijksoverheid.nl/onderwerpen/voortgezet-onderwijs/documenten/kamerstukken/2024/03/06/aanbieding-conceptkerndoelen-burgerschap-en-digitale-geletterdheid>

³² <https://www.bibliotheeknetwerk.nl/sites/default/files/documents/AI-inspiratiegids.pdf>

³³ <https://www.ai-cursus.nl/>

³⁴ Zie bijvoorbeeld ook: <https://www.ai-cursus.nl/toepassingsgebied/logistiek-en-maritiem/>

De leden van de BBB-fractie lezen: “Kan dit leiden tot verhoogde toetredingsdrempels, hetgeen potentieel belemmerend is voor een productieve marktwerking.” (blz. 5). In hoeverre ontwerpt de staatssecretaris beleid om deze potentiële belemmering te voorkomen? (9)

Generatieve AI is een systeemtechnologie die ook nieuwe uitdagingen meebrengt voor het mededingingsbeleid in de digitale economie. Dit vraagt om een dynamisch, responsief en toekomstbestendig instrumentarium. De afgelopen jaren heeft het kabinet zich ervoor ingezet om de Europese spelregels voor de allergrootste techbedrijven aan te scherpen. Zo bevat de Digital Markets Act (DMA) regels om de keuzemogelijkheden en concurrentie op platformmarkten te stimuleren en helpt de Data Act bedrijven om vrijer te switchen tussen verschillende aanbieders van cloud computing-diensten voor toegang tot de benodigde rekenkracht. Op inzet van Nederland wordt in de DMA ook rekening gehouden met het feit dat grote techbedrijven vaak een ecosysteem aan diensten aanbieden. Daarmee kunnen zij macht overhevelen van de ene naar de andere markt, bijvoorbeeld door diensten met elkaar te bundelen. De DMA biedt toezichthouders nu meer mogelijkheden om dergelijke praktijken aan te pakken en zou ook kunnen worden toegepast op diensten die gebruik maken van generatieve AI. Met deze regels wordt zo tegengegaan dat toetredingsdrempels ontstaan die een gezonde marktwerking belemmeren.

Het is echter nog te vroeg om de impact van toetredingsdrempels, schaalvoordelen en netwerkeffecten op gebied van generatieve AI goed te kwantificeren. De markt is nog in opkomst, er wordt veel geëxperimenteerd en er is momenteel sprake van stevige concurrentie tussen verschillende taalmodellen en toepassingen. Het kabinet houdt daarom alle relevante ontwikkelingen op gebied van generatieve AI nauwlettend in de gaten om te zorgen voor open en concurrerende digitale markten. Waar nodig zal de inrichting van het toezicht en interactie met het mededingingsbeleid nader worden afgestemd op die ontwikkelingen.

De leden van de BBB-fractie wijzen op de volgende passage: “De uitdaging is dan ook om een markt te creëren waarin generatieve AI-toepassingen worden aangeboden die aan alle Nederlandse en Europese waarden en wetten voldoen.” (blz. 5). Hoe en door wie worden die ‘Nederlandse en Europese waarden en wetten’ vastgesteld? (10) Is de staatssecretaris het ermee eens dat die waarden en wetten een continu veranderend proces zijn en dat het continu monitoren van de waarden en wetten nodig is om de wetten en regels omtrent generatieve AI up-to-date te houden? In hoeverre worden die waarden en wetten bepaald door de Nederlandse burger, en hoe is dit verwerkt in het beleid om de markt van generatieve AI toepassingen te reguleren? (11)

Een belangrijk onderdeel van de beleidscyclus is het monitoren en evalueren van beleid. Met de leden van de BBB-fractie ben ik het dus eens dat beleid continu aan verandering onderhevig is en dat waarden en wetten gemonitord moeten worden om de regelgeving omtrent generatieve AI up-to-date te houden. Het blijvend monitoren van het huidige wettelijk kader is dan ook één van de acties die is opgenomen in de overheidsbrede visie op generatieve AI.

Vanwege de grensoverschrijdende effecten worden veel wetten ten aanzien van het digitale domein op EU-niveau vastgesteld. Een voorbeeld hiervan is de AI-verordening, die regels bevat voor het gebruik van AI in de EU. Volgens de gewone wetgevingsprocedure in de EU komt de Europese Commissie met een wetgevingsvoorstel, waarna het Europees Parlement en de Raad van de Europese Unie over dat voorstel onderhandelen en de wet (als het voorstel wordt aangenomen) vaststellen. Omdat het Europees Parlement rechtstreeks wordt gekozen door de stemgerechtigden in de EU, is er dus een betrokkenheid van de Nederlandse burger bij het vaststellen van EU-wetgeving. In Nederland gaat dit proces, en de betrokkenheid van burgers daarbij, op een soortgelijke wijze, nu de regering wetten maakt in samenwerking met de Eerste en Tweede Kamer. Daarnaast kunnen Nederlandse burgers via internetconsultatie suggesties doen voor de verbetering van beleid en wet- en regelgeving die in voorbereiding is.

In de AI-verordening zijn verschillende mogelijkheden opgenomen om de regelgeving aan te passen wanneer dat nodig is, bijvoorbeeld via zogenaamde gedelegeerde en uitvoeringshandelingen. Bij het opstellen van uitvoeringshandelingen is Nederland, net als andere lidstaten, actief betrokken. De wijzigingen kunnen worden voorgesteld door het AI-bureau (‘AI Office’) binnen de Europese Commissie als er signalen komen dat dit nodig is. Dit kan via de toezichthouders in de lidstaten, via wetenschappers in de nog op te richten Wetenschappelijke Raad en via belanghebbenden in het Advies Forum.

Voor de regels rondom AI-modellen voor algemene doeleinden (zie ook beantwoording tweede gedeelte vraag 2 Groenlinks-PvdA-fractie) aanpassingen komen over welke informatie de aanbieder van deze modellen beschikbaar moeten houden voor toezichthouders, of op basis van welke criteria bepaald wordt of een AI-model mogelijk voor systeemrisico's zorgt.

De leden van de BBB-fractie lezen: "Dit betekent het actief betrekken van inwoners, uitvoerders, werknemers, werkgevers, ondernemers, bedrijven en professionals." Kan de staatssecretaris aangeven in hoeverre brancheorganisaties van nieuwsmedia betrokken zijn bij het vormen van een breed gedragen visie op generatieve AI? (12)

Tijdens het visietraject om te komen tot een overheidsbrede visie op generatieve AI zijn verschillende soorten sessies georganiseerd, waaronder ook in de mediasector. Zo is er in samenwerking met Electronic Commerce Platform Nederland (ECP) een sessie aanpak begeleidingsethiek georganiseerd bij RTL.³⁵ De open aanpak is ook na publicatie van de overheidsbrede visie voorgezet.

Deze leden citeren: "Verkennen we daarom onder meer manieren waarop generatieve AI kan ondersteunen om de communicatie met burgers te verhelderen en inclusiever te maken." (blz. 13). Kan de staatssecretaris specificeren wat er bedoeld wordt met inclusiever maken? Hoe wordt de communicatie inclusiever? Waarom moet communicatie inclusiever worden? (14)

Communicatie van de overheid is op dit moment niet voor iedereen begrijpelijk. U kunt bijvoorbeeld denken aan mensen die moeite hebben met lezen. Overheidsorganisaties werken aan het begrijpelijker maken van hun communicatie. De Belastingdienst is hiervan een goed voorbeeld. Zij hebben een traject waarmee zij hun 5900 standaardbrieven verbeteren. Hierin worden technische toepassingen gecombineerd met onderzoek met burgers. Met inclusiever maken van communicatie bedoel ik dat meer mensen teksten en communicatie van de overheid begrijpen.

De leden van de BBB-fractie citeren: "Ook beproeven we via pilots de meerwaarde van het toepassen van generatieve AI bij het verbeteren van overheidsdienstverlening." (blz. 13). Kan de staatssecretaris uitwerken welke pilots er lopen, deze opsommen en kort samenvatten? Hoe zijn die pilots ontworpen, welke ontwerpkeuzes liggen daaraan ten grondslag? Zijn er al resultaten van lopende pilots aanwezig? (15)

De nadere vormgeving van deze pilot in het kader van de realisatie van de visie op generatieve AI vindt momenteel nog plaats. In dit (ontwerp)traject zullen actief medeoverheden, uitvoeringsorganisaties en andere relevante stakeholders worden betrokken.

Eind 2024 zal ik uw Kamer nader informeren over de voortgang van deze, en de andere in de overheidsbrede visie opgenomen, acties.

De leden van de BBB-fractie vragen of de staatssecretaris kan aangeven welke concrete gedocumenteerde richtlijnen er zijn op het gebied van het gebruik van AI door ambtenaren. In welke ministeries zijn die richtlijnen aanwezig en zou de staatssecretaris die verschillende richtlijnen kunnen verzamelen en naar de Kamer kunnen sturen (dus niet de richtlijn vanuit de overheidsbrede visie, maar de richtlijnen op het gebied van het gebruik van AI zoals ambtenaren deze in de praktijk gebruiken)? (17)

De Kamerbrief Voorlopig standpunt voor Rijksorganisaties bij het gebruik van generatieve AI biedt algemene richtlijnen voor het wel of niet inzetten van generatieve AI.³⁶ Daarnaast ontwikkelen meerdere ministeries departementale richtlijnen voor het gebruik van generatieve AI. Bovendien wordt gewerkt aan overheidsbrede Handreiking Generatieve AI. Deze wordt na de zomer gedeeld met uw Kamer.

Verschiedende instrumenten zijn ontwikkeld om bij inzet door Rijksorganisaties de risico's van algoritmen in te schatten. Een voorbeeld hiervan is het Impact Assessment Mensenrechten en Algoritmes (IAMA). Daarnaast is het Algoritmekader in ontwikkeling. Het Algoritmekader heeft tot

³⁵ <https://begeleidingsethiek.nl/wp-content/uploads/2024/01/Verslag-workshop-Generatieve-AI-RTL.pdf>

³⁶ Kamerstuk 2023/2024, 26 643, nr. 1098.

doel overheden te ondersteunen bij de verantwoorde inzet van algoritmen. Daartoe biedt het Algoritmekader een overzicht van de belangrijkste normen en maatregelen waaraan voldaan moet worden, en normen en maatregelen die niet verplicht zijn maar als handreiking dienen voor het waarborgen van publieke waarden.

De leden van de BBB-fractie citeren: “Dit doen we vanuit een waardengedreven benadering.” (blz. 13). Kan de staatssecretaris beschrijven wat deze waarden precies zijn, en wie die heeft vastgesteld en waar hierover meer informatie te vinden is? (18)

In de Werkagenda Waardengedreven Digitaliseren³⁷ worden verschillende (publieke) waarden benoemd die van belang zijn binnen onze digitale samenleving. Dit gaat bijvoorbeeld om veiligheid, democratie, zelfbeschikking, non-discriminatie, participatie, privacy en inclusiviteit, zoals ook benoemd in het coalitieakkoord Rutte IV. In de geactualiseerde Werkagenda³⁸ – die ik eind december 2023 met uw Kamer deelde – wordt daarbij aangegeven dat het beschermen van publieke waarden als het fundament van een verantwoorde digitale transitie kan worden beschouwd. Deze waarden dienen leidend te zijn bij het ontwikkelen en gebruiken van digitale toepassingen.

De leden van de BBB-fractie lezen: “Dit vraagt om verdere inzet op het vergroten van mediawijsheid van verschillende doelgroepen.” (blz. 13). Kan de staatssecretaris aangeven of AI gelijke kansen biedt aan theoretische en praktisch geschoolde mensen, of dat daar een wezenlijk verschil tussen is, en als hier een wezenlijk verschil in zit, hoe dit voorkomen kan worden? (19)

Het is voor het kabinet van belang dat eenieder op dezelfde wijze kan profiteren van de kansen die generatieve AI te bieden heeft. Uit een recent gepubliceerd rapport³⁹ van het IMF, blijkt dat niveau van educatie mogelijk een rol speelt bij het in staat stellen tot het benutten van deze kansen. In hun onderzoek wordt gesteld dat theoretisch geschoolde mensen in hun carrière vaker dan praktisch geschoolde mensen zullen worden blootgesteld aan (generatieve) AI, en bovendien dat zij vaker banen hebben waarbij AI hen kan ondersteunen, in plaats van vervangen. Met andere woorden, theoretisch geschoolde mensen hebben een grotere kans om te werken in banen die worden beïnvloed door AI. Terwijl theoretisch geschoolde werknemers hierdoor mogelijk ook meer risico lopen op de negatieve gevolgen van AI, stelt de aard van hun werk hen tegelijkertijd beter in staat de kansen optimaal te benutten, dan dat praktisch geschoolde werknemers dat zijn.

De overheid voorkomen dat de komst van generatieve AI enige ongelijkheid tussen deze twee demografische groepen veroorzaakt en/of vergroot. Daarom zijn in de visie acties geformuleerd om dit effect te voorkomen dan wel tegen te gaan. Zo wordt er ingezet op het uitbreiden van een continue maatschappelijke dialoog met inwoners, werknemers, vakbonden en ondernemers over de impact en rol van generatieve AI op hun levens en de samenleving als geheel. De Nederlandse AI Parade van de NL-AI Coalitie heeft hierin een belangrijke rol, mogelijk ook via (verdere) uitbreiding naar o.a. het onderwijs en de Legal Parade⁴⁰. Ook worden bijvoorbeeld initiatieven gestimuleerd die de digitale vaardigheden en digitaal bewustzijn van mensen in Nederland versterken, zodat zij bewust, kritisch en actief kunnen omgaan met AI-toepassingen.

Deze leden citeren: “Een proactieve benadering is hierbij essentieel, waarbij de overheid vanaf het begin kaders biedt om de ontwikkeling van generatieve AI in goede banen te leiden en ongewenste (generatieve) AI te weren.” Kan de staatssecretaris aangeven in hoeverre experts het mogelijk vinden om ongewenste (generatieve) AI te weren? Wat is het percentage waarin dit mogelijk is? (20) Recentelijk vond er een wedstrijd op het internet plaats om een bepaalde vorm van generatieve AI ondanks veiligheidseisen toch een output te geven die zou omschreven kunnen worden als onwenselijk. Het is dus mogelijk om met zeer specifieke input te komen tot een onwenselijke output. In hoeverre is de staatssecretaris ervan op de hoogte dat ondanks hoge veiligheidseisen en waarden het mogelijk is om hiertoe te komen? Wat is volgens de

³⁷ <https://open.overheid.nl/documenten/ronl-0de79e5c4c0c9b203c0a1c263efca7eca410958b/pdf>

³⁸ <https://open.overheid.nl/documenten/8fb16ed3-0946-49d5-bf1a-96724f1762d6/file>

³⁹ Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E., & Tavares, M. M. (2024). Gen-AI: Artificial Intelligence and the Future of Work.

⁴⁰ <https://nlaic.com/bouwsteen/mensgerichte-ai/juridisch/legal-parade/>

staatssecretaris het acceptabele punt waarop het voorkomen van ongewenste uitkomsten door AI bestreden is? (21)

Het weren van ongewenste (generatieve) AI is voor het kabinet van groot belang. Het mitigeren van het risico op ongewenste AI zal bestaan uit verschillende componenten, gezien de complexiteit van de technologie. Bovendien is dit een continu proces dat samenwerking vereist tussen beleidsmakers, wetenschappers en de industrie. Experts uit deze sectoren zijn zich bewust van het probleem en werken aan methoden om de risico's te beperken, maar er is geen vastgesteld percentage waarin dit mogelijk is. Relevante methoden betreffen onder andere het testen op bias in generatieve AI-modellen.⁴¹ Het kabinet stimuleert onderzoek naar verschillende bestaande methodes en zet in op mogelijke ontwikkeling van nieuwe methodes, onder andere door middel van het Rijks AI-validatieteam. Dit team van developers buigt zich onder meer over het meetbaar maken van risico's en kansen van (generatieve) AI. Daarnaast heeft het kabinet zich in Europees en internationaal verband hiervoor ingezet, bijvoorbeeld via de AI-verordening (zie ook gedeelte van de beantwoording over o.a. de AI Office).

De leden van de BBB verwijzen naar onwenselijke output die, na het geven van zeer specifieke input, door generatieve AI kan worden gegenereerd. Het kabinet is op de hoogte van dit probleem, dat in jargon 'jailbreaking' wordt genoemd. Binnen de context van LLM's, verwijst jailbreaking naar het ontwerpen van prompts (input) met de intentie model biases uit te buiten om zo output te genereren die niet in lijn is met het doel van het model. Het model zal bijvoorbeeld antwoord geven op vragen die normaliter door het model niet zouden worden beantwoord.⁴² Vooropgesteld staat dat de mogelijkheid van jailbreaking ook voor de ontwikkelaars van generatieve AI een ongewenst resultaat is en dat actief onderzoek wordt gedaan naar verschillende mitigatiestrategieën. Desalniettemin is op het moment van schrijven geen waterdichte methode beschikbaar. Het acceptabele punt waarop het voorkomen van ongewenste uitkomsten door AI bestreden is, is dan ook niet gemakkelijk vast te stellen, aangezien niet duidelijk is hoe de technologie zich verder zal ontwikkelen. De overheid vindt het daarom van groot belang de ontwikkelingen in generatieve AI nauwgezet te volgen. Dit betekent dat beleid in nauw contact blijft met medeoverheden, uitvoeringsorganisaties, kennisinstututen, commerciële partijen, publieke belangenorganisaties en burgers. Het uitgangspunt is hierbij dat de ontwikkeling en inzet van generatieve AI moet voldoen aan bestaande wetgeving, waaronder ook de AI-verordening, en dat risico's op bijvoorbeeld jailbreaking zo goed mogelijk worden gemitigeerd met beschikbare methodes.

Daarnaast biedt de AI-verordening afdwingbare eisen voor transparante en veilige AI-modellen voor algemene doeleinden, met of zonder systeemrisico's. Generatieve AI valt hier ook onder. Nederland heeft zich tijdens de onderhandelingen actief voor deze eisen ingezet. Ontwikkelaars van AI-modellen voor algemene doeleinden zijn verplicht om technische documentatie over het model op te stellen en beschikbaar te houden voor toezichthouders en om informatie op te stellen voor ontwikkelaars die het AI-model in hun eigen AI-systeem willen integreren. Deze informatie moet inzicht geven in de capaciteiten en de beperkingen van het AI-model, waaronder bias. Als een AI-model voor algemene doeleinden ook voor systeemrisico's kan zorgen, moet de aanbieder ook een modevaluatie uitvoeren, systeemrisico's beperken, informatie over incidenten zo snel mogelijk met de toezichthouder delen en voor een passend niveau van cyberveiligheid zorgen.

Deze eisen kunnen worden afgedwongen door het AI-bureau ('AI Office'), die toezicht zal houden op de eisen aan AI-modellen voor algemene doeleinden. Zij kunnen in sommige gevallen er ook voor kiezen om een AI-model voor algemene doeleinden van de markt te halen.

Vragen en opmerkingen van de leden van de CDA-fractie en antwoord

De leden van de CDA-fractie hebben met interesse kennisgenomen van de visie en hebben hierover nog enkele vragen.

De leden van de CDA-fractie vragen of de staatssecretaris een update wil geven over de laatste stand van zaken ten aanzien van de AI-verordening. Deze leden vragen ook wanneer de staatssecretaris de Nederlandse uitvoeringswet verwacht naar de Kamer te sturen. (1)

Op 2 februari hebben alle lidstaten ingestemd met de AI-verordening en op 13 maart heeft het Europees Parlement ingestemd. In juli wordt de wet gepubliceerd en deze treedt 20 dagen na publicatie in werking.

⁴¹ <https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>

⁴² https://proceedings.neurips.cc/paper_files/paper/2023/file/fd6613131889a4b656206c50a8bd7790-Paper-Conference.pdf

Op 1 augustus 2024 worden de verschillende onderdelen van de wet gefaseerd van kracht. Dit betekent dat Nederlandse overheden en bedrijven zich in die periode kunnen voorbereiden op de eisen die voor hen gaan gelden als aanbieder of gebruikers van AI. De termijn wanneer de eisen van kracht worden is afhankelijk van het risico van de AI-systemen. Zo zijn de verboden al na 6 maanden van kracht en voor de hoog-risico AI-toepassingen gelden termijnen van 24 en 36 maanden. In die periode wordt de uitvoeringswet opgesteld in consultatie met belanghebbenden en behandeld door het parlement. Na die termijn kunnen de toezichthouders handhavend optreden als AI-systemen niet aan de eisen van de AI-verordening voldoen. Voor AI-modellen voor algemene doeleinden, waaronder de meeste grote AI-modellen die content genereren vallen, zijn de eisen na 12 maanden van toepassing en binnen die termijn wordt het Europese toezicht hierop ingericht.

Naar verwachting wordt de uitvoeringswet in 2025 naar de kamer gestuurd na het doorlopen van alle stappen bij het opstellen van een uitvoeringswet zoals de internetconsultatie en toetsing regeldruk.

De leden van de CDA-fractie vragen of de staatssecretaris de mening deelt dat de opkomst van generatieve AI het belang van digitale inclusie en digitale geletterdheid alleen maar groter maakt. En zo ja, dan vragen deze leden vragen hoe dit tot uiting komt in extra stappen op dit gebied aanvullend op de acties uit de Werkagenda. Deze leden vragen welke plaats de ontwikkeling van generatieve AI heeft binnen de aanpak van digitale inclusie en digitale geletterdheid. (2)

Deze mening deel ik. Voor de inzet vanuit de Werkagenda en nieuwe ontwikkelingen verwijs ik de CDA-fractie naar het antwoord op de vraag van de BBB-fractie over hoe burgers worden geholpen om generatieve AI te begrijpen en gebruiken. Daarnaast wijs ik de CDA-fractie op de adviesaanvraag die het kabinet heeft gedaan bij de SER over AI, de toekomst van werk en sociaaleconomische implicaties. In deze adviesaanvraag wordt expliciet aandacht gevraagd voor bij- en omscholing van werkenden in de context van generatieve AI, zowel voor ICT'ers als niet ICT'ers.

De leden van de CDA-fractie delen de mening dat sterk toezicht nodig is op de risico's van generatieve AI. Deze leden lezen dat het kabinet het van belang acht 'om te blijven evalueren of toezichthouders beschikken over de kennis en kunde, capaciteit en middelen om hun taken nu en in de toekomst effectief te kunnen uitvoeren.' Deze leden vragen hoe de staatssecretaris dit concreet aanpakt. (3)

In Nederland worden zelfstandige bestuursorganen (ZBO's) periodiek geëvalueerd om te zorgen dat ze effectief en efficiënt functioneren binnen het overheidsbeleid. Dit geldt ook voor een deel van de toezichthouders die actief zijn binnen het digitale domein (zoals bijvoorbeeld de AP, ACM en de AFM). In deze zbo-evaluatie wordt een breed scala aan aspecten bekeken, waaronder doeltreffendheid en doelmatigheid.

Sinds medio 2023 wordt de inrichting van het toezicht op de Europese AI-verordening voorbereid. Welke toezichthouders toezichthoudende bevoegdheden krijgen over de verplichtingen in de AI-verordening kan nog niet definitief worden aangegeven. Duidelijk is dat er veel verschillende toezichthouders betrokken zullen zijn vanwege het brede scala aan AI-toepassingen die onder de verordening vallen.

De AI-verordening schrijft voor dat toezichthouders op producten die al worden gereguleerd via Europees recht, zoals liften en speelgoed, ook verantwoordelijk worden voor het toezien op AI die in die producten gebruikt wordt. Daarnaast bevat de AI-verordening een aantal toepassingsgebieden waarvoor geldt dat de AI-systemen die daar worden ingezet hoog-risico zijn en aan een set eisen moet voldoen voordat ze op de markt mogen worden gebracht of worden gebruikt. Dat betreft uiteenlopende gebieden zoals kritieke infrastructuur, onderwijs, werving en selectie en rechtshandhaving. Ook worden sommige AI-systemen verboden, zoals bepaalde toepassingen van emotieherkenning in het onderwijs en op de werkvloer.

Vanwege hun expertise ligt het in de rede dat de huidige toezichthouders op die diverse aandachtsgebieden ook een rol krijgen bij het toezicht op AI in dat werkveld. De betrokken toezichthouders zijn met elkaar in gesprek om te kijken hoe het toezicht in alle gevallen het best kan worden ingericht. Actieve samenwerking is van belang, dit krijgt bijvoorbeeld vorm via de Algoritme- en AI-Kamer van het SDT, waarin afstemming plaatsvindt over toezichtsvraagstukken rondom AI. Een belangrijk aandachtspunt is dat AI-expertise veelgevraagd en dun gezaaid is. Er

wordt dus ook gekeken hoe daar mee kan worden omgegaan. Dit alles worden de komende periode verder uitgewerkt door het kabinet en de toezichthouders.

De leden van de CDA-fractie vragen naar de samenhang van deze overheidsbrede visie met de Nationale Technologie Strategie, waarin AI is aangewezen als één van de tien sleuteltechnologieën die het Nederlandse verdienvermogen moet stutten. Deze leden vragen wat de rol van generatieve AI hierin is. (4)

AI is met datatechnologie één van de 10 sleuteltechnologieën van de Nationale Technologiestrategie (NTS). De NTS is een agenda om het onderzoeks-en innovatie ecosysteem verder te versterken om in 2035 tot een sterke combinatie van AI en datatechnologie in Nederland te komen. De NTS omvat alle vormen van AI, ook generatieve AI. Naast de NTS is ook de Overheidsbrede visie op generatieve AI uitgebracht. Deze visie richt zich specifiek op grotere AI-modellen om content als tekst, beeld, geluid of computercode te genereren. De visie omvat het bredere overheidsbeleid op dit gebied en gaat in op wet- en regelgeving, internationale ontwikkelingen, toezicht en innovatie. Bij de doorontwikkeling van de NTS zullen ook inzichten op het gebied van generatieve AI worden meegenomen.

Vragen en opmerkingen van de leden van de SP-fractie en antwoord

De leden van de SP-fractie hebben kennisgenomen van de overheidsbrede visie generatieve AI, en hebben hierover een aantal vragen aan de staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties.

De leden van de SP-fractie hebben een aantal vragen over het gebruik van generatieve AI in de kunst- en cultuursector. Allereerst vragen deze leden of de staatssecretaris voldoende in beeld heeft hoe generatieve AI-modellen worden getraind op de kunst en werk van kunstenaars, auteurs, acteurs, makers van grafische media, musici, en vele andere beroepen. Kent de staatssecretaris het NRC interview met de Britse AI-hoogleraar Steward Russell op 6 nov. jl., waarin deze onder meer zegt: "Als we AI inzetten om werk te doen dat nu door mensen wordt gedaan, zal de opkomst van AI leiden tot enorme werkloosheid."⁴³ Voor de professionals in de culturele en creatieve sector is het nog schrijnender. Het werk van de mensen in die sector wordt gescrept en ingezet om generatieve AI-modellen te trainen, waardoor AI werk kan doen en maken waarmee het werk van echte en levende kunstenaars en creatieve makers overbodig en in verhouding veel duurder wordt. Hierdoor verdienen grote techbedrijven zoals Google en OpenAI veel geld aan het harde werk van kunstenaars en makers en worden diezelfde kunstenaars, naast dat hun eerdere werk grootschalig wordt gescrept door diezelfde en veel goedkoper genererende AI-modellen, ook nog eens van hun toekomstige werk en dus inkomen beroofd. De leden van deze fractie zijn van mening dat deze praktijken onrechtmatig, onrechtvaardig en vooral ook onwenselijk zijn. Is de staatssecretaris het hiermee eens? (1)

Ik heb kennisgenomen van het interview met professor Steward Russell van november vorig jaar in NRC. Op dit artikel ga ik nader in bij de beantwoording van vraag 5 van de SP-fractie.

Als een bedrijf AI-gegenereerd werk wat bedoeld is om de kunst of stijl van bepaalde personen na te bootsen gebruikt in hun advertenties en/of producten, is er dan sprake van plagiaat? Welke stappen kunnen kunstenaars en makers nu nemen om hun werk uit de trainingsdata van generatieve AI te halen, zoals een opt-out? (2)

Artiesten kunnen verschillende stappen nemen om hun werk uit de trainingsdata van generatieve AI te halen. Bij verschillende AI-bedrijven is het mogelijk om als artiest eigen werk uit de trainingsdata van AI-modellen te verwijderen. Via OpenAI is het bijvoorbeeld mogelijk eigen werk te laten verwijderen uit de trainingsdata van DALL-E, de afbeeldinggenerator van OpenAI⁴⁴. Dit kan worden bewerkstelligd via een verzoek aan OpenAI, of door de web-scraping tool van OpenAI te blokkeren voor de eigen website⁴⁵.

⁴³ AI-expert: 'Als de overheid niet snel ingrijpt, gaan we in de komende twaalf maanden de consequenties voelen'. NRC, 6 november 2023. <https://www.nrc.nl/nieuws/2023/11/06/hoogleraar-ai-als-de-overheid-niet-snel-ingrijpt-gaan-we-in-de-komende-twaalf-maanden-de-consequenties-voelen-a4179949?t=1707986947>.

⁴⁴ <https://openai.com/index/dall-e-3>

⁴⁵ https://share.hsforms.com/1_OuT5tfFSpic89PqN6r1CQ4sk30

Andere bekende afbeeldinggenerators zijn Midjourney en Stability AI, die stellen artiesten de beschermen via de Digital Millennium Copyright Act (DMCA). De DMCA is een Amerikaanse wet die ziet op de bescherming van auteursrechten in de digitale wereld. Onder de DMCA kunnen artiesten Midjourney en Stability AI verzoeken hun werk uit te sluiten van trainingsdata.

Voor Nederlandse artiesten geldt dat het auteursrecht makers in staat stelt mee te delen in de opbrengst van de exploitatie van hun werken van letterkunde, wetenschap en kunst en biedt bescherming tegen aantasting van hun werken en hun reputatie (zie ook beantwoording vraag 8 Groenlink-PvdA-fractie).

Het is voor artiesten ook mogelijk hun werk te beschermen door middel van verschillende tools, die op verschillende manieren digitale afbeeldingen bewerken zodat deze niet meer bruikbaar zijn voor het trainen van AI-modellen. Een voorbeeld hiervan is Nightshade⁴⁶.

Tot slot moet worden benadrukt dat bovenstaande methodes niet waterdicht zijn; reeds is gebleken dat het alsnog een aanzienlijke uitdaging kan zijn voor artiesten om hun werk te laten verwijderen uit trainingsdatasets.

Voor de beantwoording van het onderdeel van deze vraag dat betrekking heeft op het vraagstuk van plagiaat, verwijs ik u naar de beantwoording van vraag 8 door de fractie GroenLinks-PvdA (pagina's 5 en 6).

Is de staatssecretaris van mening dat kunstenaars en creatieve makers voldoende worden beschermd? (3)

Meer in het algemeen zijn er in Nederland regelingen die werkenden beschermen tegen tegenslagen. Mensen die in dienstverband werken, kunnen bij verlies van hun baan aanspraak maken op inkomensondersteuning via de WW. De participatiewet dient als vangnet om te voorkomen dat mensen bij verlies van inkomen onder bestaansminimum komen, ook als zij niet in dienstverband werkten maar als zelfstandige.

Deelt de staatssecretaris de mening dat er bij alle AI-gegenereerde werken duidelijk moet worden weergegeven dat het door AI is gegenereerd? (4)

De AI-verordening verplicht het om bij door AI-gegenereerde content kenbaar te maken dat deze door AI gemaakt is. Deze verplichting geldt in verschillende vormen voor zowel de ontwikkelaar van het AI-systeem dat deze content kan genereren als de gebruiker die de content daadwerkelijk genereert met het systeem.

Veel kunstenaars en makers zijn boos en bang over het gebruik van hun werk in de training van generatieve AI-modellen. Zo heeft de Kunstenbond hierover een position paper geschreven. Is de staatssecretaris bekend met deze position paper? Zou zij hierop een uitgebreide reactie kunnen geven? Is zij bereid de toegelichte punten in deze paper over te nemen?⁴⁷ (5)

Ja, ik heb kennisgenomen van deze position paper. De leden van de SP-fractie vragen terecht aandacht voor de betekenis van generatieve AI voor de kunst- en cultuursector. Mij is uiteraard bekend dat generatieve AI wordt getraind met werken van letterkunde, wetenschap en kunst en ander beschermd materiaal. Mij is ook bekend dat de voornoemde sector ervoor vreest dat de voortbrengselen van generatieve AI succesvol kunnen gaan concurreren met de prestaties waarmee die AI is getraind. Ik ken ook het interview met Stewart Russell, waarin hij waarschuwt dat AI enorme werkloosheid kan veroorzaken. Die waarschuwing verdient wat mij betreft enige nuancering. Niet alle werkzaamheden die door mensen worden verricht kunnen door AI kunnen worden vervangen. Van de werkzaamheden die wel door AI kunnen worden overgenomen, is te verwachten dat dit niet geheel maar slechts gedeeltelijk zal gebeuren. Verder zullen ontwikkelingen op het gebied van generatieve AI ook nieuwe banen creëren. Over het algemeen genomen, verwacht ik dat in ieder geval op de korte en waarschijnlijk ook de middellange termijn eerder sprake zal zijn van verandering van werkinhoud dan van totale vervanging. Dit is ook in lijn met de inbreng van een deskundig panel tijdens de presentatie van de overheidsbrede visie. Hoewel de

⁴⁶ <https://nightshade.cs.uchicago.edu/whatis.html>

⁴⁷ Position paper: generatieve kunstmatige intelligentie. Kunstenbond, 8 februari 2024. <https://kunstenbond.nl/nieuws/lees-hier-onze-position-paper-over-ai/>.

werkgelegenheid in het geheel naar verwachting niet sterk verandert, kan AI wel de positie van sommige groepen werkenden verzwakken. Sociale zekerheid biedt hen een vangnet. Op de huidige krappe arbeidsmarkt zullen zij vaak snel nieuw werk vinden, eventueel met behulp van scholing of ondersteuning bij het zoeken naar werk.

De Kunstenbond onderzoekt de mogelijkheden voor een rechtszaak tegen onrechtmatige training van AI. Het geven van een antwoord op de vraag of gelet op de uitzondering inzake tekst- en datamining sprake is van (on)rechtmatige training is aan de rechtsprekende macht voorbehouden die daarbij rekening kan houden met de specifieke omstandigheden van het geval. Zoals hiervoor op de vragen van de leden van de fracties van GroenLinks-PvdA, D66 en BBB al is geantwoord, voorziet de Auteurswet in het voetspoor van de richtlijn auteursrecht in de digitale eengemaakte markt in een uitzondering op het auteursrecht voor het maken van een reproductie van een werk van letterkunde, wetenschap of kunst ten behoeve van tekst- en datamining. Hieronder zou ook het trainen van generatieve AI kunnen ressorteren. De Kunstenbond merkt in zijn position paper terecht op dat voor alle uitzonderingen op het auteursrecht de driestappentoets geldt. Op grond van die toets gelden uitzonderingen en beperkingen alleen: (1) in bepaalde bijzondere gevallen, (2) voor zover zij niet in strijd zijn met de normale exploitatie van het beschermde materiaal en (3) de gerechtvaardigde belangen van rechthebbenden niet onevenredig worden geschaad. De driestappentoets fungeert in de eerste plaats als een instructienorm voor de wetgever die de toets bij het opstellen van elke uitzondering zorgvuldig toepast. Het resultaat van die toepassing ligt als het ware besloten in de voorwaarden die aan de inroepbaarheid van de uitzondering worden gesteld. Tegelijkertijd biedt dit kader de rechter een zekere houvast bij de eventueel verdrags- of richtlijnconforme-interpretatie van de uitzondering.

Mochten ondernemingen (waaronder Big Tech) voor de ontwikkeling van AI een geslaagd beroep op de uitzondering inzake tekst- en datamining kunnen doen, dan is het maken van een tijdelijke kopie van beschermd werk toegestaan zonder voorafgaande toestemming van rechthebbenden. Er is ook geen billijke vergoeding verschuldigd. Rechthebbenden die niet willen dat hun beschermde prestaties worden gebruikt om generatieve AI te trainen, kunnen daar zelf paal en perk aan stellen door een uitdrukkelijk voorbehoud te maken. Zo'n voorbehoud geldt louter en alleen met het oog op de toekomst. Kopieën van werken waartoe rechtmatig toegang was verkregen en die ten behoeve van tekst- en datamining zijn gemaakt voordat een voorbehoud is gemaakt, worden daardoor niet geraakt. Het is overigens technisch gezien ook niet mogelijk om kunstmatig aangeleerde intelligentie terug te draaien of af te leren. Wanneer de ontwikkelaars van generatieve AI beschermde prestaties willen gebruiken ten aanzien waarvan rechthebbenden het maken van kopieën ten behoeve van tekst- en datamining uitdrukkelijk op passende wijze hebben voorbehouden door middel van machinaal leesbare middelen bij een online ter beschikking gestelde prestaties, dan hebben zij daarvoor toestemming van de rechthebbenden nodig. Rechthebbenden kunnen aan het verlenen van toestemming voorwaarden verbinden zoals betaling van een billijke vergoeding. Op grond van de AI-verordening zal verantwoording afgelegd moeten gaan worden over met welke data (waaronder auteurs- en nabuurrechtelijk beschermde prestaties) AI is getraind. Dit dient om rechthebbenden in staat te stellen te controleren of door hen gemaakte voorbehouden en afspraken worden gerespecteerd.

Het nabootsen van de stijl van een nog in leven zijnde kunstenaar met AI levert geen auteursrechtinbreuk op, maar kan onder omstandigheden onrechtmatig zijn. De Hoge Raad heeft in 2013 overwogen dat stijl niet voor auteursrechtelijke bescherming in aanmerking komt.⁴⁸ Maar, niet uitgesloten is dat slaafse nabootsing van een stijl of van stijkenmerken onder bijkomende omstandigheden onrechtmatig kan zijn. Prof. mr. D.J.G. Visser heeft verdedigd dat de opzettelijke en herkenbare nabootsing van de persoonlijke stijl van een nog levende kunstenaar met behulp van AI onrechtmatig is. Ook wordt door hem verdedigd dat aanbieders van generatieve AI verplicht zijn deze vorm van gebruik van hun generatieve AI technisch onmogelijk te maken. Zij zijn namelijk verantwoordelijk voor het ontstaan van deze mogelijkheid, zij profiteren ervan, zij moedigen het aan, lokken het uit en zijn in de positie om dit gebruik effectief tegen te gaan. Hiervoor is geen beroep op de Auteurswet of de AI-verordening nodig. Het is gewoon onrechtmatig, aldus Visser.⁴⁹ Er

⁴⁸ HR 29 maart 2013, ECLI:NL:HR:2013:BY8661, NJ 2013/504, IER 2013/40, m.nt. F.W. Eijsvogels, AMI 2013/3, m.nt. J.C.S. Pinckaers (Broeren/Duijsens).

⁴⁹ D.J.G. Visser, Stijlnabootsing met AI is onrechtmatig, *NJB*, 2023/36, p. 3171-3175.

zijn toepassingen, zoals DALL-E, die (daarom) opdrachten om werken in de stijl van een nog levende kunstenaar na te bootsen, weigeren.⁵⁰

De Kunstenbond pleit in het hiervoor al genoemde position paper ook nog voor de introductie van een heffing. Art. 4 van de richtlijn auteursrecht in de digitale eengemaakte markt verplicht daartoe niet en er is dan ook geen enkele lidstaat die daarin voorziet. Discussie over de juridische mogelijkheid, beleidsmatige wenselijkheid en praktische uitvoerbaarheid van een dergelijk ingrijpend voorstel hoort plaats te vinden in het kader van de EU.

Ook hebben de leden van de SP-fractie vragen over het gebruik van generatieve AI bij de overheid. In de overheidsbrede visie generatieve AI lezen deze leden dat de staatssecretaris mogelijkheden ziet om met gebruik van generatieve AI processen te verbeteren, het algemeen functioneren van de overheid te verbeteren, en de dienstverlening aan burgers te optimaliseren. Zou de staatssecretaris dit kunnen toelichten, en specifiek kunnen ingaan op hoe generatieve AI kan worden gebruikt bij fraudeopsporing en dienstverlening? (6)

Zoals ook in de overheidsbrede visie op generatieve AI staat opgenomen, ziet het kabinet de mogelijkheden van generatieve AI om tot (pro-actievare) dienstverlening te komen. Generatieve AI zou bijvoorbeeld kunnen worden ingezet om (routine)taken uit te voeren, zoals het beantwoorden van veelgestelde vragen. Ook biedt generatieve AI mogelijkheden om beter in te spelen op (specifiek) taalniveau en behoeften van burgers.

Om dit in de praktijk te beproeven zullen er in 2024 pilots worden gestart waarin de meerwaarde van generatieve AI voor (proactieve) dienstverlening verder wordt verkend.

De visie bevat echter geen verwijzingen naar de potentie die generatieve AI, al dan niet, biedt voor fraudeopsporing.

De leden van de SP-fractie zijn van mening dat inzet van AI om fraude op te sporen niet moet worden toegestaan als mensenrechten niet gewaarborgd kunnen worden. Deelt de staatssecretaris deze mening? Is de staatssecretaris van plan om op grote schaal chatbots in te zetten bij de Rijksoverheid? (7)

De inzet van (generatieve) AI moet altijd in lijn met geldende wet- en regelgeving – zoals bijvoorbeeld de Grondwet en de AVG en straks de AI-verordening – gebeuren. Dit geldt zonder meer ook voor AI die wordt ontwikkeld of ingezet met het oog op de fraudeopsporing.

Verder stimuleren wij bij overheden de inzet van praktische instrumenten zoals een Impact Assessment Mensenrechten en Algoritmen (IAMA) en de handreiking non-discriminatie by design. Deze instrumenten zijn ontwikkeld om publieke waarden en mensenrechten te borgen bij de inzet van algoritmes en AI.

Zoals ook opgenomen in de overheidsbrede visie, kan generatieve AI mogelijk van meerwaarde zijn als het gaat om het verbeteren van de overheidscommunicatie met inwoners. Dit zou in sommige gevallen bijvoorbeeld kunnen gaan om chatbots. Via pilots wil de overheid dit de komende jaren (op beperkte schaal) verder beproeven. Datzelfde geldt voor het inzetten van generatieve AI voor het verbeteren van overheidsdienstverlening, bijvoorbeeld de inzet van deze technologie om burgers proactiever te helpen bij het wijzen op welke regelingen zij recht hebben.

De leden van de SP-fractie zien daarin een risico dat chatbots ter vervanging komen van menselijk contact. Is de staatssecretaris het met deze leden eens dat het altijd makkelijk moet zijn voor inwoners om menselijk contact te hebben met de overheid? (8)

Het kabinet vindt het van groot belang dat iedereen kan meedoen in de (digitale) samenleving. Dat betekent dat er ondersteuning moet zijn voor burgers met een hulpvraag en dat mensen deze ook weten te vinden. Hiervoor zijn er bijvoorbeeld circa 700 Digitale Overheid (IDO'S) versterkt.⁵¹ Bij IDO's kunnen burgers op laagdrempelige wijze (fysiek) terecht voor bij vragen over het gebruik van de digitale overheid.

⁵⁰ [OpenAI unveils DALL-E 3, allows artists to opt out of training | TechCrunch](#)

⁵¹ Zie ook: <https://www.digitaleoverheid.nl/werkagenda-waardengedreven-digitaliseren/>

Ten slotte hebben de leden van de SP-fractie vragen over het gebruik van generatieve AI in het onderwijs. De staatssecretaris geeft namelijk aan dat deze technologie leerlingen kan ondersteunen bij het maken van samenvattingen, het uitleggen van leerstof en het creëren van oefenvragen. Deze leden vragen of dit een verbetering van het onderwijs is. Met dit soort gebruik van generatieve AI worden praktijken die juist goed zijn om te oefenen overgelaten aan software. Welke impact denkt de staatssecretaris dat het gebruik van generatieve AI op deze manier heeft op de leesvaardigheid van leerlingen? (9)

Het gebruik van Generatieve AI hoeft niet ten koste te gaan van de leesvaardigheid van leerlingen, zolang leerlingen bij verschillende vakken langere uitdagende kennisrijke teksten lezen en er tijdens de les ruimte is om daarover te praten en te schrijven. (Generatieve) AI kan juist kansen bieden voor de leesvaardigheid van leerlingen, zeker als zij uitgedaagd worden om de waarde en de inhoud van de gegenereerde teksten te evalueren en erop te reflecteren.

Om een positieve impact op de leesvaardigheid te realiseren, is het belangrijk dat scholen AI doordacht inzetten. Kennisnet ondersteunt scholen hierbij, onder meer met praktische handreikingen.

Daarnaast investeren OCW en EZK met financiering van het Nationaal Groeifonds in het Nationaal Onderwijslab AI. In het lab werken onderzoekers, leraren en het bedrijfsleven samen aan de ontwikkeling van intelligente technologieën om de kwaliteit van het funderend onderwijs te verbeteren en om de pedagogische, maatschappelijke en sociale gevolgen van intelligente technologie in onderwijs inzichtelijk te maken. Één van de projecten is erop gericht om door middel van Automatic Speech Recognition het technisch leren lezen op de basisschool te verbeteren.

Ook lezen de leden van deze fractie dat docenten generatieve AI kunnen gebruiken om op basis van studentendata uit het verleden, een voorspelling te maken over het toekomstig presteren van studenten. Deze leden zien de meerwaarde die deze praktijk kan hebben. Deze data kan gebruikt worden om te zien welke studenten meer ondersteuning nodig hebben. Tegelijkertijd zijn deze leden van mening dat docenten dit sowieso moeten doen, en zij vragen of het gebruik van generatieve AI hiervoor alleen nodig is vanwege het tekort aan leerkrachten. Is deze werkwijze niets meer dan een pleister plakken op een diepe wond? (10)

Generatieve AI kan leraren en docenten ondersteunen bij de complexiteit van hun vak. Bijvoorbeeld door bepaalde onderwijstaken over te nemen, te vergemakkelijken of te versnellen,⁵² zoals het genereren van nieuw lesmateriaal of om onderwijsmethodes te personaliseren. Hierdoor behoudt de docent meer tijd voor bijvoorbeeld de begeleiding van leerlingen en studenten. Het is daarom van belang dat het onderwijs de kansen van digitalisering benut, zodat de kwaliteit van het onderwijs toeneemt. Het is onduidelijk of de inzet van digitalisering zoals generatieve AI leidt tot lagere werkdruk bij docenten.⁵³

Het lerarentekort is een urgent probleem en het kabinet neemt maatregelen om deze tekorten aan te pakken. De ministers van OCW hebben in de voortgangsrapportage eind 2023 de stand van zaken weergegeven van de lerarenstrategie.⁵⁴

Vragen en opmerkingen van de leden van de Volt-fractie en antwoord

De leden van de Volt-fractie hebben kennisgenomen van de Overheidsbrede visie op generatieve AI (hierna: “visie”) van het kabinet. De Volt-fractie wil het kabinet en in het bijzonder de staatssecretaris voor digitalisering complimenteren met de uitgebreide visie. In de visie en de bijbehorende stukken blijkt dat het kabinet de ontwikkelingen op het gebied van AI nauwgezet volgt en blijkt een grote ambitie.

Over de visie hebben de leden van de Volt-fractie nog enkele vragen die zien op de existentiële risico's van AI, de internationale samenwerking en eerdere moties die de Volt-fractie heeft ingediend.

⁵² <https://www.onderwijsraad.nl/publicaties/adviezen/2022/09/28/inzet-van-intelligente-technologie>

⁵³ <https://www.rathenau.nl/nl/digitalisering/naar-hoogwaardig-digitaal-onderwijs>

⁵⁴ Kamerstuk 2023–2024, 27 923, nr. 460.

De leden van de Volt-fractie zijn groot voorstander van internationale en Europese samenwerking op het gebied van AI-regulering en beleid. In dat licht vragen de leden in hoeverre het kabinet heeft gekeken naar beleid van andere (gelijkgestemde) landen op het gebied van generatieve AI en welke elementen uit dat beleid zij hebben overgenomen en of er elementen zijn waar zij expliciet is afgeweken van standaarden uit die landen? (1)

De leden vragen ook op welke punten samenwerking met gelijkgestemde landen nog niet voldoende functioneert en wat het kabinet daar aan doet. (2)

Generatieve AI is een grensoverschrijdend fenomeen. Voor de totstandkoming van de visie zijn internationale ontwikkelingen op het gebied van generatieve AI dan ook actief gemonitord. Daarmee zijn verschillende (generatieve) AI-strategieën geanalyseerd, zoals de AI-strategie uit het VK⁵⁵. Hierin wordt onder andere het belang van investeren in fysieke infrastructuur en het monitoren van AI-ontwikkelingen genoemd, wat in onze visie ook wordt benadrukt.

Daarbij komt dat er verschillende internationale gremia bestaan waar kennisontwikkeling met betrekking tot AI plaatsvindt. Nederland speelt een actieve rol in bijvoorbeeld het Comité voor AI (CAI) van de Raad van Europa en de AI-expertgroep van de OESO. Andere gremia die worden gevolgd betreffen onder andere de VN High Level Advisory Body on AI en het G7 Hiroshima AI-proces. Daarnaast is er tijdens de totstandkoming van de AI-verordening nauw samengewerkt met de lidstaten uit de Europese Unie.

Verschiedende bevindingen uit bovenstaande strategieën en gremia tonen gelijkenissen met de visie. Hierbij moet worden gesteld dat Nederland zich onderscheidt als een van de eerste EU-lidstaten met een visie die specifiek focust op generatieve AI. Deze nadruk op het gebied van generatieve AI plaatst Nederland in de voorhoede van generatieve AI governance binnen de Europese Unie.

Kinderen en jongeren hebben breed toegang tot generatieve AI. Soms zelf zonder dat zij zich daarvan bewust zijn. Die toegang heeft grote impact op hen en zij kunnen zich daar minder makkelijk tegen weren dan de rest van de samenleving. In dat licht hebben de leden van de Volt-fractie een aantal vragen: hoe wordt geborgd dat de stem van kinderen en jongeren mee wordt genomen bij het ontwikkelen van beleid en wetgeving aangaande (generatieve) AI? (3) Kan het kabinet toelichten waarom in het hoofdstuk over de wettelijke kaders, het Internationaal Verdrag inzake de Rechten van het Kind niet is opgenomen? In het geval het kabinet dit wenselijk vindt, is zij bereid om dit wettelijk kader toe te voegen aan de visie? (4)

Generatieve AI staat als onderwerp geagendeerd voor de Jongerenraad Digitalisering. Deze Jongerenraad, die ik samen met UNICEF oprichtte, adviseert mij vier keer per jaar over digitaliseringsvraagstukken en biedt dus de uitgelezen mogelijkheid om de stem van kinderen en jongeren mee te nemen bij de ontwikkeling van wetgeving en beleid, óók aangaande (generatieve) AI. Verder merk ik op dat het in de overheidsbrede visie benoemde juridische kader niet uitputtend is en dat het Internationaal Verdrag inzake de Rechten van het Kind uiteraard ook van toepassing is. Het kabinet heeft veel aandacht voor de rechten van kinderen en jongeren, óók online. De acties die hieraan bijdragen zijn uitgebreid uiteengezet in de Kamerbrief kinderrechten en digitalisering van juni 2023 (Kamerstukken II 2022-2023, 26643, nr. 1063).

De leden van de Volt-fractie hebben meerdere moties op het gebied van AI ingediend in de afgelopen jaren. Nu de AI-verordening een definitieve tekst heeft, vragen deze leden of het kabinet al werk maakt van het opnemen van een wettelijke verplichting voor overheden om algoritmes te registreren in een register.⁵⁶ (5)

Het ministerie van BZK onderzoekt momenteel de mogelijkheid om het algoritmeregister een verplicht karakter te geven en in hoeverre de AI-verordening hier al in voorziet en of er dan nationaal nog aanvullende verplichtingen nodig zijn. Nu de teksten van de AI-verordening bekend zijn kan een start worden gemaakt om te bezien hoe de wetgeving rondom het algoritmeregister kan worden vormgegeven. Het ministerie betreft daarbij ook de ervaringen met de publicatie van

⁵⁵ <https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version>

⁵⁶ Kamerstukken II 2021/22, 35925-VII-26 (motie van het lid Dassen c.s. over het gebruik van het algoritmeregister voor overheden verplicht stellen), via: [Vaststelling van de begrotingsstaten van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties \(VII\) voor het jaar 2022 | Tweede Kamer der Staten-Generaal](#).

algoritmes tot dusverre. Daar is een publicatiestandaard gemaakt dat een voorlopig onderscheid maakt tussen drie classificaties: niet alleen hoog risico, maar ook impactvolle AI en/of algoritmes.

Met betrekking tot de motie Dassen om een impactanalyse per ministerie uit te voeren vragen deze leden het kabinet of zij van mening is dat de motie hierbij volledig is uitgevoerd. Zo ja, waarom? Zo niet, hoe wordt verder invulling gegeven aan de motie?⁵⁷ (6)

Zoals ook aangegeven in de Verzamelbrief van 9 april jl., is de desbetreffende motie van het lid Dassen vanwege de brede scope afgedaan via verschillende wegen.⁵⁸ De overheidsbrede visie op generatieve AI bevat een uitgebreide uiteenzetting van de (potentiële) impact van generatieve AI op een breed scala aan beleidsterreinen. Daarnaast biedt ook het Rathenau-rapport over generatieve AI, dat in opdracht van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) is uitgevoerd, een overzicht van kansen en risico's van deze technologie op verschillende domeinen van de samenleving.⁵⁹ Ook heeft het kabinet de SER in januari jl. gevraagd om een advies op te stellen over de kansen van AI om arbeidsproductiviteit te verhogen, de kansen en risico's in kwaliteit van werk, en de impact op werkgelegenheid en inkomens(on)gelijkheid.⁶⁰

Een risico dat de leden van de Volt-fractie nog onderbelicht vinden in de visie is het mogelijke existentiële risico dat AI met zich meebrengt.⁶¹ Hoewel de leden inzien dat AI grote kansen meebrengt en er aandacht moet zijn voor de fundamentele problemen in het hier en nu, zijn zij van mening dat de overheid een verantwoordelijkheid heeft om alle risico's af te wegen. In dat licht vragen deze leden het kabinet om op de existentiële risico's te reflecteren en om dit in de visie te verwerken. (7)

Het kabinet en de medeoverheden hebben er in de overheidsbrede visie voor gekozen om te kijken naar een breed pakket aan kansen en risico's. Het eerste uitgangspunt van de visie is het veilig ontwikkelen en toepassen van generatieve AI. Hierbij hoort een degelijke en wetenschappelijke inschatting van welke risico's er zijn en hoe deze het beste aangepakt kunnen worden.

De Leden van Volt vragen specifiek naar de existentiële risico's die door AI veroorzaakt zouden kunnen worden. Uit een recente uitvraag onder meer dan 2700 AI-onderzoekers blijkt dat een meerderheid de kans op menselijk uitsterven groter inschat dan 5 procent.⁶² Dit onderzoek bevestigt de oproep die we recent hebben gezien in verschillende brieven, ondertekend door wetenschappers en AI-onderzoekers over de dreigingen van AI.

Over de precieze kans dat deze dreiging daadwerkelijk een realiteit wordt, en de termijn waarop dit soort risico's op zouden kunnen gaan optreden, is weinig overeenstemming. Zoals de vraagstellers ook al aangeven, sluit aandacht voor deze risico's niet uit dat we ook aandacht hebben voor risico's die momenteel al effect hebben.

Het kabinet erkent op lange termijn de dreigingen van AI-systemen. De beste manier om deze aan te pakken is door ze op te nemen als een van de risico's die worden aangepakt in ons beleid, zonder dat dit de aandacht voor andere risico's uitsluit. We bereiden ons voor door niet alleen te focussen op de huidige uitdagingen, maar ook proactief op toekomstige risico's. De beleidsoplossingen voor de verschillende vormen van risico's komen vaak overeen, zoals het doen van onderzoek naar het verbeteren van de transparantie en veiligheid van modellen, het komen tot internationale afspraken en standaarden, en het monitoren van de capaciteiten van AI-systemen. Dit zijn acties die het kabinet al onderneemt en al een plek hebben in de visie.

Daarnaast is recent de AI-verordening officieel aangekomen. Twee belangrijke pijlers van de AI-verordening zijn⁶³:

⁵⁷ Kamerstukken II 2023/24, 36 410, nr. 53.

⁵⁸ Kamerstuk 2023/24, 26 643, nr. 1149.

⁵⁹ Zie ook: Rathenau Instituut (2023), Rathenau scan over generatieve AI.

⁶⁰ Zie ook: <https://www.rijksoverheid.nl/documenten/rapporten/2024/01/18/adviesaanvraag-ser-ai-algoritmes-en-data-de-toekomst-van-werk-en-sociaaleconomische-implicaties>

⁶¹ NRC 19 juli 2023, [Opinie | Existentieel risico of niet, we moeten AI beheersen - NRC](#).

⁶² https://aiimpacts.org/wp-content/uploads/2023/04/Thousands_of_AI_authors_on_the_future_of_AI.pdf

⁶³ [Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world - Consilium \(europa.eu\)](#)

- *Regels voor AI-modellen voor algemene doeleinden die systeemrisico's kunnen veroorzaken, evenals voor AI-systemen met een hoog risico. Volgens de AI-verordening moeten aanbieders van AI-modellen voor algemene doeleinden met systeemrisico's een grondige modevaluatie uitvoeren en mogelijke systeemrisico's beperken. Onderdeel van de modevaluatie is ook het adversarial testen van het model. Dit is een evaluatiemethode die in een veilige omgeving de ingebouwde veiligheidsmechanismen van een AI-model uittest en probeert te omzeilen. Op deze manier kan inzichtelijk worden gemaakt hoe veilig een model is in de echte wereld. Het in kaart brengen en mitigeren van mogelijke systeemrisico's betekent natuurlijk ook dat mogelijke existentiële risico's nadrukkelijk afgewogen moeten worden. Zodra er serieuze incidenten plaatsvinden, is de aanbieder van het AI-model voor algemene doeleinden met systeemrisico's ook verplicht om het Europese AI-bureau ('AI Office') meteen in te lichten. Het AI-bureau kan ook uit eigen beweging onderzoek doen en maatregelen nemen, waaronder het model van de markt halen.*
- *Betere bescherming van (mensen) rechten door de aanbieders van hoogrisico AI-systemen een verplichting op te leggen om een "fundamental rights impact assessment" uit te voeren voordat het wordt ingezet.*

Hierbij wel de nuance dat de AI-verordening niet van toepassing is op systemen die uitsluitend voor militaire of defensiedoeleinden worden gebruikt. Op dezelfde manier bepaalt het akkoord dat de verordening niet van toepassing zou zijn op AI-systemen die uitsluitend worden gebruikt voor onderzoek en innovatie, of voor mensen die AI gebruiken om niet-professionele redenen.

Naast de overheidsbrede visie op generatieve AI blijkt deze benadering ook uit de inzet van Nederland op andere thema's rondom AI, zoals het inzetten op het verantwoordelijke gebruik van AI in het militaire domein. In lijn met deze visie is het kabinet toegewijd aan het continu evalueren en bijstellen van ons beleid om zowel de huidige als toekomstige risico's te mitigeren, gesteund door actueel wetenschappelijk inzicht en technologische ontwikkelingen. Dit komt ook terug in de lerende aanpak die gekozen is in de visie.

Uit de Scan Generatieve AI van het Rathenau Instituut blijkt dat het huidige beleid de risico's op het gebied van non-discriminatie, mededinging, arbeidsomstandigheden, veiligheid en desinformatie onvoldoende adresseert. Daarnaast concludeert het Rathenau Instituut dat generatieve AI-systemen invloed hebben op de sociale, cognitieve en culturele ontwikkeling van mensen. De leden van de Volt-fractie vragen het kabinet in hoeverre zij zich in deze conclusies herkent en of het kabinet zich verantwoordelijk en geschikt acht om deze risico's te verminderen en welke maatregelen zij voorstelt om deze risico's te verminderen. (8)

Allereerst zou ik een kanttekening willen plaatsen bij (een deel van) deze vraagstelling. Het Rathenau Instituut concludeert namelijk niet dat het huidige beleid de risico's op het gebied van non-discriminatie, mededinging, arbeidsomstandigheden, veiligheid en desinformatie onvoldoende adresseert. Het Rathenau Instituut stelt daarentegen dat het een "reële mogelijkheid [is] dat het bestaande en aangekondigde beleid onvoldoende zijn opgewassen tegen deze risico's. Actie is daarom nodig. De overheid en politici moeten toetsen waar beleid aanscherping behoeft."⁶⁴ In de overheidsbrede visie generatieve AI is een veelheid aan bestaand beleid benoemd en nieuwe actielijnen aangekondigd die zien op de risico's in de genoemde domeinen. Daarin vormt de AI-verordening een belangrijke pijler, maar ook additionele acties zien op het mitigeren van de risico's. Om te toetsen of beleid adequaat is dan wel aanscherping behoeft neemt het verschillende acties, zoals ook aangekondigd in de actielijn 'nauwgezet volgen van alle ontwikkelingen' in de overheidsbrede visie op generatieve AI. Daarin wordt onder andere de Directie Coördinatie Algoritmes van AP en het AI-bureau genoemd, de EU-toezichthouder die door de Europese Commissie is ingericht. Relevant ook is de toegezegde overheidsbrede monitor van de initiatieven, ontwikkelingen en gebruik op het gebied van generatieve AI door overheden en (semi-)publieke organisaties.

Om bias en discriminatie tegen te gaan wordt daarnaast ingezet op verschillende additionele maatregelen die bijdragen aan het mitigeren van bias in generatieve-AI-systemen. Het Rijks-AI-validatieteam faciliteert publiek beschikbare benchmarking en tooling (zoals bias-detectie, op basis van bijvoorbeeld democratische input) om vangrails aan te brengen voor verantwoorde generatieve AI in Nederland.

⁶⁴ Rathenau-scan Generatieve AI, p. 42.

Wat betreft mededinging is allereerst belangrijk te noemen dat het kabinet de impact van technologiebedrijven op de Nederlandse samenleving herkent. De technologie die ten grondslag van deze bedrijven ligt heeft de potentie om de manier waarop Nederland communiceert, leert en werkt te veranderen. Generatieve AI is nog volop in ontwikkeling maar het is aannemelijk dat deze technologie bestaande dynamieken in de markt versterkt. Zo kunnen schaalvoordelen en netwerkeffecten nog belangrijker worden. Dit vereist extra aandacht vanuit de toezichthouder en wetgever. De afgelopen jaren is er verschillende Europese wetgeving voor de digitale economie tot stand gekomen met het doel mededinging in digitale markten te bevorderen. Zo beoogt de Digital Markets Act (DMA) de macht van zogeheten poortwachters te verminderen door het stimuleren van concurrentie en de keuzevrijheid voor consumenten en ondernemers te vergroten. Daarnaast maakt de Dataverordening het makkelijker voor gebruikers van clouddiensten om over te stappen tussen verschillende aanbieders. Ook draagt de Dataverordening eraan bij dat gebruikers van producten meer controle krijgen over de gegevens uit die producten zodat gebruikers ook zelf kunnen profiteren van de waarde van deze gegevens en de concurrentie in zogeheten aftermarketdiensten wordt vergroot. Ik zie dit als belangrijke stappen om goedwerkende digitale markten te realiseren. Aangezien de DMA pas sinds kort van toepassing is en de Dataverordening pas vanaf september 2025 van toepassing wordt is het nog te vroeg om de impact van de nieuwe Europese wetgeving te beoordelen. Daarnaast is de markt voor generatieve AI nog in opkomst, er wordt veel geëxperimenteerd en er is momenteel sprake van stevige concurrentie tussen verschillende taalmodellen en toepassingen. Het kabinet heeft tot op heden daarom geen aanvullende maatregelen voorgesteld om de marktdynamiek te beïnvloeden, maar houdt ontwikkelingen op het gebied van generatieve AI nauwlettend in de gaten. Waar nodig zal de inrichting van het toezicht en interactie met het mededingingsbeleid nader worden afgestemd op die ontwikkelingen. Om marktmacht van niet-Europese technologiebedrijven te beperken is het belangrijk dat er in Nederland en Europa een klimaat wordt gestimuleerd waarin er breed de ruimte is voor het experimenteren, testen en opschalen van betrouwbare en transparante (generatieve) AI-modellen en tools. Voorbeelden daarvan zijn de in de visie genoemde ALT-EDIC en GPT-NL (het Nederlandse taalmodel ontwikkeld door TNO, NFI en SURF).

Als het gaat om veiligheid is van belang te noemen dat er in Europees en internationaal opzicht veel en op verschillende wijzen wordt samengewerkt met gelijkgestemde landen op dit thema, zoals benoemd in de overheidsbrede visie onder de noemer 'internationale ontwikkelingen' en de actielijn 'samenwerken'. Binnen Nederland zetten het NCSC en de AIVD op het vergroten van de technische kennis op het onderwerp AI en weerbaarheid tegen cyberdreigingen in dit nieuwe domein en weerbaarheid tegen ongewenst gebruik van LLM's. Daarbij worden organisaties geïnformeerd over urgente ontwikkelingen en krijgen zij concrete handvatten aangeboden voor de veilige ontwikkeling van AI. Verder, om risicovolle strategische afhankelijkheden te voorkomen of te mitigeren en het geopolitieke handelingsvermogen van de EU kunnen vergroten heeft het kabinet in oktober 2023 de Agenda Digitale Open Strategische Autonomie (DOSA) gepubliceerd, waarbij AI één van de beleidsprioriteiten is.

Wat betreft arbeidsomstandigheden heeft het kabinet de Sociaal-Economische Raad gevraagd om de impact van AI (waaronder generatieve AI) op arbeidsproductiviteit en kwaliteit van werk in kaart te brengen - en advies gevraagd hoe de kwaliteit van werk (waaronder arbeidsomstandigheden) geborgd kan worden bij de steeds verdere maatschappelijke adoptie van (generatieve) AI.

Als het gaat om desinformatie, herken ik de conclusies gesteld door het Rathenau Instituut dat generatieve AI nieuwe risico's kan stellen in de creatie en verspreiding van online desinformatie. Zoals het Rathenau Instituut aangeeft in de Scan Generatieve AI, biedt de Wet inzake Digitale Diensten (DSA) juridische kaders om desinformatie aan te pakken. Zo zijn de zeer grote online platforms sinds augustus 2023 verplicht om systeemrisico's in kaart brengen die voortvloeien uit hun diensten. Desinformatie kan zo'n systeemrisico vormen. Tegen deze risico's moeten ze maatregelen nemen. Daarnaast hebben de minister van BZK en ik in november 2023 in de Kamerbrief Weerbaarheid verkiezingsproces Tweede Kamerverkiezing aangegeven dat er momenteel wordt onderzocht of onze beleidsinzet op desinformatie met oog op de nieuwe technologieën, zoals generatieve AI, moet worden aangepast. Hierover, en over de verdere invulling van maatregelen om desinformatie voortvloeiend uit AI tegen te gaan, wordt uw Kamer geïnformeerd in de Voortgangsbrief over de Rijksbrede strategie voor de effectieve aanpak van desinformatie.

Omdat de ontwikkelingen op het gebied van generatieve AI ontzettend snel gaan, is het van belang dat er ook snel ingegrepen kan worden op het moment dat er ontwikkelingen zijn die te grote risico's meebrengen. In die zin begrijpen de leden van de Volt-fractie de wens om na te denken

over een 'Rapid Response Team', maar willen zij het kabinet wel vragen om nader toe te lichten wat 'ingrijpen' precies inhoudt en hoe dit zich verhoudt tot het huidige stelsel van toezicht en handhaving. Wordt een dergelijk team bijvoorbeeld wettelijk verankerd? (9)

Een adviesorgaan heeft geen bevoegdheden die leiden tot daadwerkelijk 'ingrijpen'. Wel is het mogelijk om een adviesorgaan een adviserende rol te geven ten aanzien van beleid. Bij het instellen van een adviesorgaan dat het kabinet adviseert over beleid, is het belangrijk om te bekijken in hoeverre de Kaderwet Adviescolleges in acht moet worden genomen. In dat geval kan het zijn dat een dergelijk orgaan inderdaad wettelijk wordt verankerd. De wensen met betrekking tot en de mogelijkheden van een dergelijk adviesorgaan worden momenteel verkend.