Tweede Kamer der Staten-Generaal

2

Vergaderjaar 2023–2024

Aanhangsel van de Handelingen

Vragen gesteld door de leden der Kamer, met de daarop door de regering gegeven antwoorden

2482

Vragen van het lid **Rajkowski** (VVD) aan de Minister van Economische Zaken en Klimaat en de Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties over *het bericht dat werknemers in een brief stellen dat OpenAI roekeloos is en dat het ontbreekt aan toezicht* (ingezonden 10 juni 2024).

Antwoord van Minister **Beljaarts** (Economische Zaken), mede namens de Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties (ontvangen 9 september 2024). Zie ook Aanhangsel Handelingen, vergaderjaar 2023–2024, nr. 2097.

Vraag 1
Bent u bekend met het bericht dat werknemers in een brief stellen dat OpenAI roekeloos is en dat het ontbreekt aan toezicht?[1] Herkent u de zorgen zoals geuit in de brief van huidig en voormalig werknemers van OpenAI? Zo ja, hoe ziet de Minister het mogelijke gevaar van het versterken van ongelijkheid en het verspreiden van desinformatie door AI (in Nederland)? Welke gevaren ziet de Minister nog meer?

Antwoord 1
Ja, hier ben ik mee bekend. Ik herken de risico's die worden genoemd in de brief en de bezorgdheid over het effect dat artificiële intelligentie (AI) kan hebben op het creëren en verspreiden van mis- en desinformatie. Generatieve AI zoals de producten van OpenAI maken het in potentie makkelijker om mis- en desinformatie te creëren. Om deze risico's te beperken zijn regels opgesteld in de AI-verordening. Zo moeten aanbieders van AI-systemen die synthetische content[2] genereren ervoor zorgen dat de output van het AI-systeem wordt gemarkeerd in een machineleesbaar formaat en detecteerbaar zijn als kunstmatig gegenereerd of gemanipuleerd. Dit maakt het makkelijker voor grote online platforms om systeemrisico's die kunnen voortvloeien uit de verspreiding van kunstmatig gegenereerde of gemanipuleerde content te identificeren en te beperken. Verder geldt dat AI-systemen die bedoeld zijn om te worden gebruikt voor het beïnvloeden

[1] NOS (4 juni 2024). Werknemers in brief: «OpenAI is roekeloos en het ontbreekt aan toezicht». https://nos.nl/artikel/2523179-werknemers-in-brief-openai-is-roekeloos-en-het-ontbreekt-aan-toezicht

[2] Synthetische content is content die gemaakt of bewerkt is met behulp van AI. Dat kan o.a. gaan om foto's, video's, geluid en tekst.

ah-tk-20232024-2482
ISSN 0921 - 7398
's-Gravenhage 2024

van de uitslag van een verkiezing of referendum, of van het stemgedrag bij verkiezingen of referenda als hoog-risico worden geclassificeerd vanuit de AI-verordening. Aanbieders van deze systemen moeten dan ook voldoen aan de eisen en verplichtingen voor AI-systemen met een hoog risico, zoals het identificeren en voorkomen van risico's. OpenAI verbiedt dan ook het gebruik van hun producten voor politieke beïnvloeding.
Het kabinet wil dat deze risico's van generatieve AI worden ingeperkt en aangepakt, maar ook de kansen van generatieve AI benutten voor het tegengaan van desinformatie. In de Voortgangsbrief Rijksbrede strategie voor de effectieve aanpak van desinformatie[3], die onlangs vanuit mijn ambtsvoorgangers, aan uw Kamer is verzonden, wordt dieper ingegaan op de impact van generatieve AI en het verspreiden van desinformatie.
Het kabinet erkent ook het risico dat AI kan hebben op groeiende ongelijkheid en de sociaaleconomische implicaties. Daarom steunt zij ook het eerder aangevraagd advies van het vorige kabinet hierover aan de Sociaal Economische Raad[4], zodat het kabinet samen met werkgevers en werknemers actie kan ondernemen om de impact van AI op de samenleving op een verantwoorde manier een vervolg te geven. De AI-verordening heeft als doel om risico's van AI-systemen aan te pakken, waaronder het risico op discriminatie wat ongelijkheid veroorzaakt en kan vergroten. De AI-verordening bepaalt dat onder meer risico's op aantasting van fundamentele rechten zoals het recht op non-discriminatie moeten worden aangepakt bij de ontwikkeling van AI-systemen en bij de inzet van AI-systemen door overheidsinstanties.

Vraag 2 en 4
Hoe ziet u in algemene zin het spanningsveld tussen het feit dat de AI Act pas in haar volledige vorm over enkele jaren in werking treedt en dat het op dit moment wenselijk kan zijn om toezicht te houden en daar waar nodig in te grijpen bij organisatie die maatschappij veranderende technieken uitbrengen, zoals kunstmatige intelligentie en large language models?
Wat kan er nu al gedaan worden om te voorkomen dat (de toepassing van) generatieve en specifieke AI gevaarlijk wordt? Deelt de Minister de mening dat het belangrijk is om niet te wachten op de AI-Act en nu al stappen te ondernemen, gezien de snelle ontwikkelingen als het gaat om (de toepassing van) verschillende soorten AI? Welke maatregelen neemt de Minister hiervoor? Zo nee, waarom niet?

Antwoord 2 en 4
Als het gaat om het reguleren van AI als product binnen de EU en in Nederland, vormt de AI-verordening hiervoor het juridisch fundament. De AI-verordening is in werking getreden op 1 augustus jl. en wordt stapsgewijs van toepassing. Zo zijn bepaalde AI-praktijken al na 6 maanden verboden. De eisen voor AI-modellen voor algemene doeleinden (waar Large Language Models zoals GPT van OpenAI naar verwachting ook onder vallen) zijn al na 12 maanden van toepassing. Hieronder vallen ook de zwaardere eisen wanneer deze modellen ook voor systeemrisico's kunnen zorgen. Voor de eisen aan hoog-risico AI-toepassingseisen gelden termijnen van 24 en 36 maanden. Het is belangrijk dat ontwikkelaars en gebruiksverantwoordelijken van deze AI-modellen en systemen voldoende tijd hebben om zich goed te kunnen voorbereiden op het van toepassing worden van de eisen.
Het kabinet neemt nu al stappen om risico's van (generatieve) AI te adresseren. Daarom wordt in Nederland al hard gewerkt aan het versterken van het toezicht op AI, mede in voorbereiding op de AI-verordening. Zo is onder de Autoriteit Persoonsgegevens (AP) de Directie Coördinatie Algoritmes (DCA) opgericht om het toezicht op algoritmes en AI in Nederland te versterken. Dit doet zij door risico's van algoritmes en AI te signaleren en analyseren, samenwerking tussen toezichthouders te versterken en guidance te bevorderen. Daarnaast hebben de Rijksinspectie Digitale Infrastructuur (RDI) en de DCA in 2024 samen 3,1 miljoen euro ontvangen van het Ministerie van Economische Zaken om Nederland te kunnen voorbereiden op het inrichten

[3] https://www.rijksoverheid.nl/documenten/kamerstukken/2024/06/17/tk-voortgangsbrief-rijksbrede-strategie-voor-de-effectieve-aanpak-van-desinformatie-en-aankondiging-nieuwe-acties
[4] https://www.rijksoverheid.nl/documenten/rapporten/2024/01/18/adviesaanvraag-ser-ai-algoritmes-en-data-de-toekomst-van-werk-en-sociaaleconomische-implicaties

van toezicht op de AI-verordening. De RDI en DCA bieden in de loop van dit jaar een definitief advies aan over de inrichting van het toezicht op de AI-verordening aan de bewindspersonen van Economische Zaken, Binnenlandse Zaken en Koninkrijksrelaties en Justitie en Veiligheid. Daarna wordt de voorgenomen inrichting van het toezicht opgenomen in de uitvoeringswet die aan het parlement wordt voorgelegd via de gebruikelijke wegen.
Onder de AI-verordening zijn aanbieders van AI-modellen voor algemene doeleinden met ingang van 1 augustus 2025 verplicht om informatie en documentatie op te stellen, up-to-date te houden en beschikbaar te stellen voor toezichthouders en aanbieders die dat AI-model in hun AI-systemen willen integreren. De bedoeling is dat deze informatie en documentatie inzicht geeft in de capaciteiten en beperkingen van het AI-model voor algemene doeleinden en daarmee aanbieders verderop in de waardeketen die met dit model verder bouwen aan specifiekere AI-modellen of -systemen in staat stellen aan hun verplichtingen uit de AI-verordening te voldoen.
Als AI-modellen voor algemene doeleinden voor systeemrisico's kunnen zorgen, dan moeten ze volgens de AI-verordening aan extra eisen voldoen. Dit kan het geval zijn als er zeer veel computerkracht gebruikt is bij het trainen van het model, maar ook het aantal eindgebruikers kan een relevante factor zijn bij de beoordeling of een AI-model voor systeemrisico's kan zorgen. Modellen met systeemrisico's moeten onder andere modelevaluaties uitvoeren, risico's beoordelen en beperken en zorgen voor een passend niveau van cybersecurity. Ook moeten incidenten zo snel mogelijk gemeld worden bij het Europese AI-bureau, dat toezicht houdt op deze modellen.
Het AI-bureau («AI Office») van de Europese Commissie is bevoegd om toezicht te houden op AI-modellen voor algemene doeleinden. De eisen voor deze modellen treden 1 augustus 2025 in werking, waarna het AI-bureau kan gaan controleren of deze worden nageleefd.
Er wordt niet gewacht tot de AI-verordening van toepassing is om veilige ontwikkeling en gebruik van AI te bevorderen. Zo heeft de Europese Commissie recent het «AI Pact» gelanceerd. Dit AI Pact heeft als doel om organisaties vroegtijdig aan de AI-verordening te laten voldoen. Aan de ene kant door informatie en best practices tussen organisaties uit te wisselen, en aan de andere kant door organisaties te motiveren om toe te zeggen dat ze vroegtijdig aan de eisen zullen voldoen. Daarnaast ontwikkelt het AI-bureau al verschillende instrumenten die bijdragen aan een effectieve uitvoering van de AI-verordening, zoals richtsnoeren en lagere wetgeving die de AI-verordening verder duidelijk maken, praktijkcodes, benchmarks en het ondersteunen en bundelen van kennis uit de verschillende adviesorganen. Daarbij draagt de Commissie namens de Europese lidstaten bij aan internationale gedragscodes voor geavanceerde AI-systemen, zoals de Hiroshima gedragscode van de G7.[5] Nederland volgt deze ontwikkelingen nauw en werkt tijdens de implementatie van de AI-verordening via verschillende adviesorganen nauw samen met het Europese AI-bureau.
Daarnaast bestaan er al verschillende wetten die niet specifiek over AI gaan, maar die wel normen bevatten waarmee de inzet van AI wordt gereguleerd. De Algemene Verordening Persoonsgegevens (AVG) biedt bijvoorbeeld juridische kaders bij geautomatiseerde besluitvorming met behulp van AI en bij het verwerken van persoonsgegevens voor het trainen of gebruik van AI. Op grond van de AVG hebben EU privacy toezichthouders ervoor gezorgd dat ChatGPT extra privacy waarborgen biedt. De Algemene wet bestuursrecht (Awb) stelt normen aan de kwaliteit van overheidsbesluiten en de motivering daarvan, ook als daar AI voor wordt gebruikt. De Algemene wet gelijke behandeling (Awgb) beschermt tegen discriminatie als AI-systemen worden gebruikt bij het aanbieden van bijvoorbeeld werk, goederen en diensten.
Tenslotte zijn er de afgelopen jaren verschillende (niet-regulerende) beleidsacties op nationaal niveau opgezet die ook bijdragen aan een verantwoorde inzet van AI. Voorbeelden hiervan zijn het algoritmeregister voor AI dat wordt gebruikt door de overheid[6], de ELSA-labs[7], het investeringsprogramma

[5] Het G7 Hiroshima AI-proces richt zich op het opstellen van internationale richtsnoeren voor organisaties die geavanceerde AI-systemen ontwikkelen, en hebben tot doel veilige, beveiligde en betrouwbare AI wereldwijd te bevorderen.
[6] https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/algoritmes/algoritmeregister/

AiNed, het ROBUST programma en het impact assessment mensenrechten algoritmes (IAMA).

Vraag 3
Welke mogelijkheden zijn er in afwachting van de volledige inwerkingtreding van de AI-Act om de negatieve risico's van AI op de samenleving te mitigeren? Zijn er mogelijkheden om op te treden tegen activiteiten en handelingen die nu formeel nog niet als illegaal bestempeld zijn onder de AI-Act, maar dat wel gaan zijn zodra de volledige AI-Act in werking zal treden?

Antwoord 3
Het is pas mogelijk om juridisch op te treden tegen activiteiten en handelingen op het moment dat zij illegaal zijn. Alhoewel de AI-verordening nog niet (volledig) van toepassing is, kunnen bepaalde activiteiten en handelingen op dit moment al wel illegaal zijn op grond van bestaande juridische kaders, zoals non-discriminatiewetgeving. Bijvoorbeeld als deze activiteiten en handelingen onrechtmatig zijn.
Overigens is het denkbaar dat activiteiten en handelingen illegaal zijn, terwijl het toezicht daarop en de handhaving daarvan op dat moment nog niet zijn gerealiseerd. Zo zijn de verboden in de AI-verordening van toepassing met ingang van 1 februari 2025, terwijl het toezicht op en de handhaving van de verboden pas 6 maanden later geregeld hoeft te zijn.[8] Voor het inzetten van toezichtsbevoegdheden of handhavend optreden is een grondslag in de nationale (formele) wet noodzakelijk.[9] Tot die tijd is een overtreding van de verboden wel onrechtmatig en kan een procedure gestart worden bij de civiele rechter. Tevens kan een toezichthouder, na de aanwijzing bij of krachtens formele wet, zo nodig met terugwerkende kracht handhaven.

Vraag 5
Deelt u de mening dat de AI-adviesraad, zoals om gevraagd in de motie-Rajkowski (Kamerstuk 21 501-33, nr. 1041), er zo snel mogelijk moet komen om na te kunnen denken over mogelijke acties als het gaat om (de toepassing van) AI op het gebied van veiligheid, weerbaarheid, wenselijkheid, innovatie en strategische autonomie? Op welke termijn kunnen we een dergelijke raad verwachten? Welke maatregelen of besluiten moeten er nog genomen worden?

Antwoord 5
Genoemde motie vraagt de regering om te onderzoeken of een Nederlands adviserend orgaan van toegevoegde waarde kan zijn om de overheid op korte termijn te adviseren bij ontwikkelingen rondom kunstmatige intelligentie. Door het vorige kabinet is gemeld dat er een verkenning wordt gedaan naar het inrichten van een AI-adviesraad en de laatste stand van zaken van vóór de zomer[10]. Die verkenning continueert het huidig kabinet. 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. Ik verwacht dit najaar de uitkomsten van dit onderzoek met uw Kamer te kunnen delen.

Vraag 6
Deelt u de mening dat het wenselijk is om in te zetten op Nederlandse/Europese ontwikkeling van kunstmatige intelligentie met de focus op publieke waarden met positieve maatschappelijke en economische effecten en dat een AI-fabriek hier een interessant instrument voor kan zijn? Deelt u de mening dat het wenselijk kan zijn om een AI-fabriek, eventueel in Europese samenwerking, te plaatsen in Nederland en het daarmee een mooie aanvulling kan zijn op ons digitale knooppunt en het versterken van onze digitale economie van de toekomst? Zo ja, op welke manier gaat u zich hiervoor inzetten in Europa? Zo nee, waarom niet?

---

[7] Dit zijn labs waarin door wetenschappers, ondernemers en publieke instellingen onderzoek wordt gedaan naar de ethische, juridische en sociale aspecten van AI.
[8] Zie artikel 113 van de AI-verordening.
[9] Zie onder meer Aanwijzing 5.9 van de Aanwijzingen voor de regelgeving.
[10] Kamerstuk 26 643, nr. 1197.

Antwoord 6
Het AI-beleid in Nederland heeft zich sinds 2019 gericht op het versterken van het waardengedreven AI-ecosysteem waar onze publieke belangen, zoals fundamentele rechten, zijn geborgd en we de maatschappelijke en economische kansen verzilveren. Dit gebeurt onder meer via de publiek-private Nederlandse AI Coalitie (NLAIC), het AiNed investeringsprogramma (Nationaal Groeifonds) en de inzet voor de AI-verordening. Om het Europese AI-ecosysteem verder te versterken, heeft de Europese Commissie in januari 2024 het AI-innovatiepakket gelanceerd. In de BNC-fiches hierover is positief gereageerd op het voorstel van de Commissie om AI-fabrieken te ontwikkelen.[11] Het voorstel voor amendering van de verordening omtrent de voortzetting van de gemeenschappelijke onderneming voor High Performance Computing (HPC) is 23 mei jl. goedgekeurd door de Raad van Concurrentievermogen.
In de Kamerbrief «Verkenning mogelijkheden AI-faciliteit»[12] van 4 juni jl. is uw Kamer onlangs geïnformeerd over drie scenario's die de Ministeries van OCW, BZK en EZ momenteel aan het verkennen zijn, namelijk: 1) bestaande middelen gebruiken, 2) meer investeringen in Europese AI-faciliteiten binnen EuroHPC, en 3) AI-faciliteit in Nederland, waarbij deze faciliteit onderdeel uitmaakt van het bredere Europese supercomputerecosysteem (gemeenschappelijke onderneming EuroHPC). Per scenario brengen de betrokken ministeries momenteel de meerwaarde en haalbaarheid in kaart. Hierin is onder meer aandacht voor de impact op maatschappelijke en economische belangen.
De scenario-aanpak is bedoeld om een zorgvuldige afweging te kunnen maken door het kabinet, omdat er aanzienlijke investeringen mee gemoeid kunnen zijn.

Vraag 7
Bent u bereid om, eventueel met Europese collega's in gesprek te gaan met de Europese Commissie over hoe er op korte termijn gezorgd kan worden dat OpenAI openheid geeft over een aantal zaken zodat we het tegengaan van misbruik van kunstmatige intelligentie en andere onwenselijke effecten kunnen mitigeren? Zo nee, waarom niet?

Antwoord 7
Zoals ook aangegeven in het antwoord op vraag 2 tot en met 4, zal het AI-bureau van de Europese Commissie toezicht gaan houden op AI-modellen voor algemene doeleinden (waar Large Language Models zoals GPT van OpenAI ook onder vallen). Om ervoor te zorgen dat de aanbieders van deze modellen weten hoe ze aan de gestelde eisen moeten voldoen, worden er *codes of practice* opgesteld. Dit zijn praktische richtsnoeren die in samenwerking met de lidstaten en andere belanghebbenden worden opgesteld.
Hoewel de Commissie na een jaar bevoegdheden heeft om naleving van de regels te eisen, treedt zij in het kader van het opstellen van de *codes of practice* daarvoor al wel in dialoog met onder andere aanbieders en gebruikers van AI-modellen voor algemene doeleinden. Indien de Commissie signalen krijgt over de negatieve gevolgen van deze AI-modellen, kan hierover ook de dialoog aangegaan worden met de ontwikkelaars.
Via de recent opgerichte AI-Board is Nederland actief betrokken bij deze ontwikkelingen en staan we in nauw contact met de Europese Commissie over de bredere uitwerking van de AI-verordening.

[11] BNC-fiche «Mededeling stimuleren van startups en innovatie in betrouwbare AI» (Kamerstuk 22 112, nr. 3908) en «Fiche: Verordening supercomputerinitiatief kunstmatige intelligentie» (Kamerstuk 22 112, nr. 3907).
[12] Kamerstuk 26 643, nr. 1180