Tweede Kamer der Staten-Generaal

2

Vergaderjaar 2023–2024

26 643 Informatie- en communicatietechnologie (ICT)

26 362 Modernisering van de overheid

Nr. 1098

BRIEF VAN DE STAATSSECRETARIS VAN BINNENLANDSE ZAKEN EN KONINKRIJKSRELATIES

Aan de Voorzitter van de Tweede Kamer der Staten-Generaal

Den Haag, 11 december 2023

Tijdens het Commissiedebat Digitaliserende Overheid van 23 maart jl. heb ik uw Kamer toegezegd een quickscan uit te voeren naar de impact van generatieve AI-technologie op IT- en overheidspersoneel. Het Lid Van Weerdenburg (PVV) vroeg mij om uit te zoeken of het inzetten van generatieve AI een deel van de oplossing kan zijn om de krapte in deze sectoren minder nijpend te maken. Deze vraag raakt aan de breed gesteunde motie Dekker-Abdulaziz en Rajkowski[1] die oproept om tot een integrale visie te komen over de inzet, de potentie en de risico's van nieuwe AI-producten, zoals generatieve AI. Deze overheidsbrede visie op generatieve AI, die ik in Q1 2024 naar uw Kamer zal versturen, schetst een integraal beeld van de impact die deze technologie heeft op onze maatschappij en de uitdagingen en mogelijkheden die hiermee zijn gemoeid. Daarbij worden concrete acties aangekondigd om als samenleving de vruchten optimaal te kunnen plukken van het potentieel dat generatieve AI biedt.

Generatieve AI is een veelbelovende en kansrijke ontwikkeling om efficiënter te werken of effectiever publieke doelen te behalen. De keerzijde is dat naast verschillende kansen ook risico's bestaan in de toepassing ervan. Hierbij speelt mee dat de technologie zich zo snel ontwikkelt, dat de gevolgen niet altijd tijdig zijn te overzien. Door Rijksorganisaties is mede daarom de behoefte gedeeld om een Rijksbreed standpunt in te nemen over het gebruik van generatieve AI.
Middels deze brief ga ik in op de uitgevoerde snelle doorlichting (quickscan) en informeer ik u over mijn voorlopige standpunt voor Rijksorganisaties bij het gebruik van generatieve AI. Hierna volgt een nadere toelichting op dit standpunt. Hiervoor heb ik onder andere advies ingewonnen bij de Landsadvocaat en de Autoriteit Persoonsgegevens (AP). Afsluitend geef ik aan welke vervolgstappen ik neem in opmaat naar de overheidsbrede visie op generatieve AI.

[1] Kamerstuk 26 643, nr. 1003

kst-26643-1098
ISSN 0921 - 7371
's-Gravenhage 2023

Snelle doorlichting impact generatieve AI op (IT-) overheidspersoneel: potentie groot en behoefte aan richtlijnen

In antwoord op de vraag van het Lid van Weerdenburg (PVV) om uit te zoeken of het inzetten van generatieve AI een deel van de oplossing kan zijn om de krapte in deze sectoren minder nijpend te maken, liet ik een snelle doorlichting uitvoeren. Deze doorlichting is niet allesomvattend, maar dient als richtinggevende indicatie om verder onderzoek op voort te bouwen.

Uit de snelle doorlichting naar de impact van generatieve AI op (IT-) overheidspersoneel, die medio 2023 is uitgevoerd bij acht Rijksorganisaties, waaronder het Ministerie van Justitie en Veiligheid, het Ministerie van Defensie, de Belastingdienst en Rijkswaterstaat, is naar voren gekomen dat veel ambtenaren momenteel al experimenteren met generatieve AI-toepassingen, zoals *ChatGPT*. Ook blijkt dat generatieve AI als een kansrijke ontwikkeling wordt gezien om onder andere de arbeidsproductiviteit van (IT-)overheidspersoneel en daarmee ook de kwaliteit van de dienstverlening van de overheid te verbeteren. Mogelijke voorbeelden zijn het faciliteren van meer overheidscontactpunten met langere openingstijden, het snel genereren van content en het effectiever inzetten van ondersteunend werk. Er worden echter ook risico's genoemd. De risico's bevinden zich onder andere op het gebied van privacy, gegevensbescherming en auteursrecht, maar ook klimaat, inclusie en mensenrechten worden genoemd. Departementen en overige Rijksorganisaties geven aan behoefte te hebben aan richtlijnen voor het veilig en verantwoord gebruik van deze technologie[2].

Op basis van deze snelle doorlichting kunnen we nog geen definitieve uitspraken doen of generatieve AI een bijdrage gaat leveren om de krapte op de arbeidsmarkt minder nijpend te maken. Wel is de technologie veelbelovend en bestaat de kans dat het werk ten minste verandert. Nader onderzoek is nodig om deze kansen te kunnen verzilveren.

De vraag gesteld door lid Van Weerdenburg (PVV) acht ik hiermee beantwoord. In de bijlage vindt u een nadere uitwerking van de snelle doorlichting. In navolging van de snelle doorlichting komt er een Rijksbrede inventarisatie van de initiatieven en ontwikkelingen op het gebied van generatieve AI (zie vervolgstap 3).

Voorlopig standpunt omgang generatieve AI door Rijksorganisaties

Als overheid kiezen we voor de verantwoorde, waardengedreven inzet van AI. Dat houdt in dat dit op een zorgvuldige en veilige wijze gebeurt. Wanneer Rijksorganisaties gebruik maken van generatieve AI, is het bovendien altijd van belang dat wordt voldaan aan de geldende wet- en regelgeving. Het uitgangspunt van dit voorlopig standpunt is daarom het voldoen aan bestaande wet- en regelgeving[3]. Hiervoor heb ik onder andere advies ingewonnen bij de Landsadvocaat en de Autoriteit Persoonsgegevens (AP).

---

[2] Gezien de vrijwillige deelname en het beperkt aantal deelnemende organisaties bieden de uitkomsten van de snelle doorlichting wellicht geen compleet beeld voor de gehele rijksoverheid. In een vervolginventarisatie wordt ernaar gestreefd om dit aantal organisaties uit te breiden.

[3] Zoals de Algemene verordening gegevensbescherming, het Auteursrecht, de Wet justitiële en strafvorderlijke gegevens, de algemene beginselen van behoorlijk bestuur en de grondrechten

Mijn voorlopig standpunt voor het gebruik van generatieve AI door Rijksorganisaties is als volgt:

- Als overheid stimuleren wij innovatie en zien wij het belang in van experimenten om generatieve AI in te zetten voor publieke waarden;
- Hierbij dienen alle generatieve AI-toepassingen te voldoen aan geldende wet- en regelgeving;
- Om vast te stellen welke specifieke vorm van inzet van generatieve AI wel of niet mogelijk is, dient voorafgaand aan het gebruik ervan per unieke casus een risicoanalyse te worden uitgevoerd. Dit zijn een (pre-scan) Data Protection Impact Assessment (DPIA[4]) en een algoritme impact assessment (zoals een Impact Assessment Mensenrechten en Algoritmes (IAMA[5])), waarin de risico's en risicobeperkende maatregelen worden vastgesteld;
- De uitkomsten hiervan dienen voorafgaand aan de inzet van de toepassing ter advies aan de (departementale) Chief Information Officer en de Functionaris Gegevensbescherming te worden voorgelegd. De CIO Rijk destilleert Rijksbrede lessen en voorkomt doublering van risicoanalyses;
- De bovengenoemde punten zijn ook van toepassing bij het gebruiken of (door)ontwikkelen van een *open source* generatieve AI-toepassing. In het kader van de Wet Open Overheid en het stimuleren van transparantie, heeft *open source* generatieve AI de voorkeur;
- Niet-gecontracteerde generatieve AI-toepassingen[6] zoals *ChatGPT, Bard en Midjourney*, voldoen over het algemeen niet aantoonbaar aan de geldende privacy- en auteursrechtelijke wetgeving[7]. Zodoende is het gebruik hiervan door Rijksorganisaties (of in opdracht daarvan) in beginsel niet toegestaan, in die gevallen waarin het risico bestaat dat wetgeving wordt overtreden, tenzij de aanbieder en de gebruiker aantoonbaar voldoen aan de geldende wet- en regelgeving;
- Gecontracteerde generatieve AI-toepassingen dienen bovendien te voldoen aan de Algemene Rijksvoorwaarden bij IT-overeenkomsten 2022[8] en aan departementale inkoopvoorwaarden indien deze prevaleren; en
- Nadat aan bovenstaande punten is voldaan, is bij het gebruik van een generatieve AI-toepassing van belang dat medewerkers voldoende worden geïnformeerd over hoe zij deze technologie op een verantwoorde wijze kunnen inzetten. Dit kan door training of richtlijnen voor verantwoord gebruik (zie vervolgstap 6 en 7).

Dit voorlopige standpunt is geen categorisch verbod van de technologie, maar herhaalt geldende wet- en regelgeving. Het gebruik wordt niet ontzegd, maar gekaderd. Zo blijft het mogelijk te experimenteren met de technologie. Binnen pilots (zie vervolgstap 4), maar ook op Rijksbreed niveau wordt een diepgaander ethisch gesprek over proportionaliteit gevoerd en wordt de maatschappelijke waarde van de technologie, in samenhang met wet- en regelgeving, nader onderzocht. Dit standpunt heeft de status voorlopig, omdat de afstemming nog plaatsvindt in lijn met de overheidsbrede visie op generatieve AI. Na deze afstemming betreft het een volgende versie van het standpunt. Gezien de snelheid van de technologische ontwikkeling is een definitief standpunt pas aan de orde nadat kennis en ervaring is opgedaan met verantwoorde inzet na bredere ethische beraadslaging. Het standpunt wordt verder uitgewerkt in

---

[4] Data protection impact assessment (DPIA) | Autoriteit Persoonsgegevens

[5] Impact Assessment Mensenrechten en Algoritmes | Rapport | Rijksoverheid.nl

[6] Bijvoorbeeld openbaar toegankelijke, door grote techbedrijven ontwikkelde (en vaak online op het internet) aangeboden vormen van generatieve AI

[7] De reden dat deze bullet «in beginsel» bevat is omdat er aanbieders kunnen bestaan van niet-gecontracteerde AI-toepassingen die kunnen voldoen aan wet- en regelgeving

[8] Wetten.nl – Regeling – Besluit vaststelling Algemene Rijksvoorwaarden bij IT-overeenkomsten 2022 (ARBIT-2022) – BWBR0047124 (overheid.nl)

de Handreiking voor Rijksorganisaties bij het gebruik van generatieve AI (zie vervolgstap 6). In een vervolgversie van dit standpunt in 2024 verwacht ik meer duidelijkheid te kunnen verschaffen over in welke toepassingsgebieden generatieve AI verantwoord kan worden ingezet. Hier is meer onderzoek voor nodig en dit blijft in dit voorlopige standpunt buiten beschouwing.

Toelichting op voorlopig standpunt

Wat is generatieve AI?

Generatieve AI is een vorm van kunstmatige intelligentie die in staat is om tekst, audio, afbeeldingen, computercode en video's te creëren. In tegenstelling tot taakspecifieke AI die zich beperkt tot *analyse* van beschikbare data, richt generatieve AI zich op het *creëren van nieuwe resultaten* op basis van al beschikbare data. Daarbij is de kwaliteit van de output zodanig, dat het soms lastig te onderscheiden is van door mensen gemaakte content. Eén van de meest herkenbare toepassingen van generatieve AI voor het brede publiek zijn AI-chatbots. Deze digitale assistenten kunnen communiceren op een manier die sterk lijkt op menselijke interactie. Een bekend voorbeeld hiervan is *ChatGPT*, een chatbot op basis van een *large language model* (LLM). Dit soort modellen is gespecialiseerd in natuurlijke taalverwerking en richt zich op het voorspellen en verwerken van tekst. Hoewel deze modellen in de basis zijn ontworpen voor taalbewerking, ligt hun kracht en groeiend succes in de veelzijdige toepassingen ervan, variërend van het programmeren van code tot het spelen van bordspellen.
Een belangrijk onderdeel van generatieve AI is de data waarmee het model getraind wordt. Het trainen van deze modellen vraagt om ongekende hoeveelheden data, bijvoorbeeld afkomstig van het internet. Wanneer niet duidelijk is op welke data deze modellen zijn gebaseerd, kunnen vragen op het vlak van onder andere privacy en auteursrecht niet beantwoord worden en kunnen inbreuken niet worden uitgesloten. Ook heeft dit effect op de output van de modellen. Zo kan de gebruikte data leiden tot bias en onjuistheden in die output.

Verschil tussen niet-gecontracteerde en gecontracteerde generatieve AI

Wanneer ik kijk naar de juridische implicaties van het gebruik van generatieve AI, maak ik onderscheid tussen twee soorten toepassingen: *niet-gecontracteerde generatieve AI-toepassingen* en *gecontracteerde generatieve AI-toepassingen.* Dit onderscheid is nodig, omdat de contracten met de aanbieders van generatieve AI sterk van invloed zijn op de wijze waarop het mogelijk is om ongewenste risico's te beperken. Wanneer we als overheid gebruik maken van niet-gecontracteerde diensten betekent dit dat we geen nadere afspraken kunnen maken over de verwerking van ingevoerde data, en geen garanties kunnen krijgen over de herkomst van trainingsdata. Dat brengt ongewenste gevolgen met zich mee. Wanneer afspraken niet mogelijk zijn, kan de gebruiker er alleen maar voor kiezen om de toepassing onder de gestelde algemene voorwaarden te gebruiken, of om de toepassing helemaal niet te gebruiken.
Ik versta onder niet-gecontracteerde generatieve AI de openbaar toegankelijke, door derden ontwikkelde en (vaak online op het internet) aangeboden vormen van generatieve AI, zoals *ChatGPT* voor tekst of *Midjourney* voor afbeeldingen, inclusief de onderliggende *foundation models*. Deze vormen van generatieve AI onderscheid ik van gecontracteerde en ingekochte zakelijke varianten en eventuele door de rijksoverheid zelf ontwikkelde modellen, waarbij zowel de trainingsdata als het gebruik aan voorwaarden kunnen worden verbonden. In tegenstelling tot

niet-gecontracteerde toepassingen, kunnen daarbij wel degelijk wederkerige afspraken worden gemaakt over het gebruik en de ontwikkeling van een toepassing.
Het onderscheid tussen niet-gecontracteerde en gecontracteerde generatieve AI is echter niet allesomvattend. Zo kunnen *open source* generatieve AI-toepassingen bijvoorbeeld buiten deze categorisering vallen. Bij het nagaan of het gebruik van een specifieke toepassing is toegestaan, dient daarom altijd te worden gehandeld in de geest van het voorlopige standpunt, door een risicoanalyse uit te voeren en na te gaan of er wordt voldaan aan wet- en regelgeving.

Waarvoor is dit standpunt bedoeld?

Dit standpunt geldt zowel bij het uitvoeren van bedrijfsvoeringtaken als primaire publieke taken. Het standpunt is ook van toepassing wanneer een externe partij diensten of producten levert aan een Rijksorganisatie met gebruikmaking van generatieve AI. Ook strekt dit standpunt zich uit tot het experimenteel gebruik van generatieve AI in bijvoorbeeld pilots of proeftuinen.

Voor wie is dit standpunt bedoeld?

De doelgroepen van dit voorlopige standpunt zijn departementen en Rijksorganisaties[9]. In de komende maanden wil ik in gesprek gaan met medeoverheden om een overheidsbreed standpunt in te nemen.

Hoe is dit voorlopig standpunt tot stand gekomen?

Op basis van de interne richtlijnen van de Europese Commissie[10] is interdepartementaal gewerkt aan een concept standpunt. Hier hebben wij in lijn met de interne richtlijnen van de Europese Commissie in eerste instantie gefocust op «online beschikbare» generatieve AI vanwege de grote risico's die hierbij gelden. Inmiddels bezigen wij de term «online beschikbaar» niet meer en geldt het voorlopig standpunt voor alle generatieve AI. Na het concept standpunt gedeeld te hebben met de Functionarissen Gegevensbescherming (FG's), is ten eerste door de departementale Functionarissen Gegevensbescherming tezamen geadviseerd om inzichtelijk te maken in hoeverre de werkterm «online beschikbare» generatieve AI voldoet aan de eisen van de AVG. Ten tweede is aan de Landsadvocaat de vraag voorgelegd welke juridische mogelijkheden er bestaan voor de inzet van «online beschikbare» generatieve AI gezien huidige, geldende wetgeving. Ten slotte is een advies gevraagd aan de Autoriteit Persoonsgegevens die ook hierop heeft geadviseerd. In de bijlage van deze brief vindt u de volledige adviezen, die hieronder nader worden toegelicht.
De specifieke bedenkingen in de adviezen hebben betrekking op de volgende punten:

- De trainingsdata, veelal verkregen via grootschalige *scraping* (het extraheren van informatie van webpagina's) van openbare bronnen op internet of andere digitale bronnen, kunnen privacygevoelige gegevens bevatten;
- Het getrainde generatieve AI model kan onjuiste of bevooroordeelde gegevens over personen bevatten;

---

[9] Begrensd door wettelijke of gerechtvaardigde uitzonderingen die van toepassing zijn bij belangenafweging in het kader van bijvoorbeeld nationale veiligheid, opsporing, rechtshandhaving, Defensie of inlichtingenverzameling.

[10] Europese Commissie (2023, 24 mei). Guidelines for staff on the use of online available generative artificial intelligence tools. (COM-AI-GUIDELINES.pdf (politico.eu)).

- De interactievragen van de gebruikers («prompts») kunnen worden hergebruikt om het model te *finetunen*, waardoor onbedoeld gevoelige informatie openbaar kan worden gemaakt;
- Een generatieve AI-toepassing kan zeer gevoelige informatie afleiden uit de interactie met de gebruiker;
- De rechten van betrokkenen zoals de mogelijkheid tot inzage- en het eventueel verwijderen, of rectificeren van (onjuiste) persoonsgegevens zijn momenteel zeer beperkt;
- Het gebruik van generatieve AI brengt auteursrechtelijke risico's met zich mee, omdat onduidelijk is of aanbieders voldoende rekening houden met de rechten van auteurs, wiens werken zij bij de ontwikkeling van hun modellen gebruiken. Veel aanbieders zijn momenteel niet transparant over welke auteursrechtelijke werken zij gebruiken in het trainingsproces; en
- Vanuit de rechtspraak is er nog geen expliciet oordeel over of, en in welke vorm, *scraping* -ten behoeve van het trainen van generatieve AI modellen- auteursrechtelijk gezien toelaatbaar is.

Vervolgstappen

De rijksoverheid wil het potentieel van kansrijke technologieën benutten en het gebruik van generatieve AI binnen de rijksoverheid op een veilige en verantwoorde manier bevorderen. Daarom neemt het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties het voortouw om de volgende acties uit te voeren:

1. Het bovenstaande standpunt wordt helder gecommuniceerd binnen de organisaties van de rijksoverheid;
2. Het bovenstaande standpunt wordt gedeeld met medeoverheden en de vervolgstappen worden in samenwerking vormgegeven;
3. Er komt er een inventarisatie van de initiatieven en ontwikkelingen op het gebied van generatieve AI binnen de rijksoverheid;
4. Binnen verschillende Rijksorganisaties worden pilotprogramma's in veilige en gecontroleerde omgevingen uitgevoerd. Na een Rijksbrede inventarisatie zal het Ministerie van Binnenlandse Zaken aanhaken op deze pilotprogramma's om inzicht te krijgen in de specifieke toepassingsgebieden. Hiervoor wordt een community opgezet zodat overheidsbreed van elkaar geleerd kan worden en snel kan worden opgeschakeld indien een generatieve AI-toepassing na risicoanalyses wordt geaccepteerd;
5. Met aanbieders gaan we om de tafel om de mogelijkheden te onderzoeken om individueel gebruik van generatieve AI in de huidige zakelijke werkomgeving verantwoord aan te bieden;
6. Er wordt een «Handreiking voor Rijksorganisaties bij het gebruik van generatieve AI» geschreven om generatieve AI op een verantwoorde wijze te laten inzetten door Rijksorganisaties;
7. Door middel van trainingen (via de RijksAcademie voor Digitalisering en Informatisering Overheid) en andere bijeenkomsten (via de Generatieve AI community) wordt kennisdeling gefaciliteerd over de mogelijkheden voor veilig gebruik van generatieve AI;
8. De interbestuurlijke inkoopvoorwaarden worden aangescherpt met het oog op generatieve AI, waarbij publieke waarden als veiligheid, transparantie, non-discriminatie, privacy- en gegevensbescherming en duurzaamheid worden geborgd; en
9. We nemen actief deel aan de totstandkoming van de Europese AI-verordening en het AI-Verdrag van de Raad van Europa. De Europese wet stelt eisen stellen aan de ontwikkeling en het gebruik van generatieve AI, die in Nederland bindend zullen zijn en derhalve deel uitmaken van het Rijksbrede standpunt. De volgende versie van

dit standpunt is onderdeel van de overheidsbrede visie op generatieve AI, die Q1 2024 met uw Kamer wordt gedeeld. Eind 2024 vindt een herijking van het standpunt plaats.

De Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties,
A.C. van Huffelen