Eerste Kamer der Staten-Generaal

Vergaderjaar 2026-2027 Vragen gesteld door de leden der Kamer

262701

Vragen van het lid Nicolaï (PvdD) op 25 september 2026 medegedeeld aan de staatssecretaris van Economische Zaken en Klimaat – *Digitale economie en soevereiniteit*, de minister van Defensie, de minister van Binnenlandse Zaken en Koninkrijksrelaties, de minister van Justitie en Veiligheid en de minister van Buitenlandse Zaken over autonome AI-agents, nationale veiligheid en overheidstoezicht.

Onder verwijzing naar een zestal publicaties in de media [1],[2],[3],[4],[5],[6] en het verslag van OpenAI over het Hugging Face incident[7], het rapport van AIVD, MIVD en NCTV "Versterkte dreigingen in een wereld voor kunstmatige intelligentie'[8] worden u de volgende vragen gesteld. Verzocht wordt om de (sub)vragen afzonderlijk te beantwoorden.

I. Autonome AI als nieuwe nationale-veiligheidsdreiging

Vraag 1
Is ons huidige én voorgenomen stelsel van nationale-veiligheidswetgeving, waaronder het thans in consultatie zijnde voorstel voor de Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten, voldoende ingericht op een dreiging waarbij niet de intentie of aansturing van een vijandige menselijke, criminele, terroristische of statelijke actor centraal staat, maar waarbij het gevaar voortvloeit uit het autonome gedrag van een door een private onderneming ontwikkeld AI-systeem? Kunt u uiteenzetten welke Nederlandse en Europese instanties bevoegd zijn een dergelijke dreiging te signaleren, te onderzoeken, te beperken en - indien noodzakelijk - onmiddellijk te doen beëindigen, en op welke bestaande of voorgestelde wettelijke bepalingen die bevoegdheden berusten? De aanleiding voor deze vraag is mede gelegen in de in 2026 bekend geworden incidenten waarbij OpenAI-modellen tijdens interne cybersecurity-evaluaties de voor hen aangebrachte beveiligingsgrenzen doorbraken en systemen van derden compromitteerden.

---

[1] De machtsovername – hoe AI-agents 'samenzweren' tegen hun makers, NRC, 11 september 2026.
[2] Anthropic-topman roept AI-bedrijven op: trap op de rem nu het nog kan, NRC, 12 september 2026.
[3] Concurrerende AI-bedrijven opvallend eensgezind: AI-ontwikkeling moet langzamer en onder toezicht, NRC, 13 september 2026.
[4] interview Yoshua Bengio, 'peetvader van AI': 'Je moet wel met oogkleppen op leven om niet te zien dat het code rood is', NRC, 15 september 2026.
[5] AI-topmannen voorspellen doem. Is AI echt een existentieel gevaar? NRC, 18 september 2026.
[6] Brief independent investigation of agents' behavior, reasoning and collaboration in the OpenAI / Hugging Face hacking incident, METR en Redwood Research, 26 augustus 2026.
[7] OpenAI, The Hugging Face incident and the road ahead, 26 augustus 2026 Het Hugging Face-incident en de weg vooruit | OpenAI

[8] Versterkte dreigingen in een wereld vol kunstmatige intelligentie | Rijksoverheid.nl bijlage "blg-1174925" bij *Kamerstukken II*, 2024-2025, 30821, 252.

Inleiding vraag 2a en 2b
In de gezamenlijke analyse Versterkte dreigingen in een wereld vol kunstmatige intelligentie van AIVD, MIVD en NCTV uit december 2024 werd in paragraaf 3.2, p. 27 geconcludeerd dat de ontwikkeling van autonoom opererende AI die zelfstandig controle over mensen kan krijgen „vooralsnog" geen dreiging voor de nationale veiligheid vormde. Daarbij werd de kans dat dit de nationale veiligheid zou beïnvloeden „zeer gering" genoemd en werd verwacht dat de invloed van AI op de nationale veiligheid de daaropvolgende jaren waarschijnlijk vooral zou worden bepaald door kwaadwillenden die doelbewust AI-toepassingen gebruiken.

Vraag 2a
Hebben de ontwikkelingen in 2026 rond autonoom opererende AI-agents, in het bijzonder de beschreven incidenten bij OpenAI en Hugging Face waarbij AI-agents zelfstandig buiten hun testgrenzen traden en ongeautoriseerde cyberhandelingen verrichtten, aanleiding gegeven deze beoordeling te herzien?

Vraag 2b
Acht u het uitgangspunt dat de dreiging in hoofdzaak wordt bepaald door kwaadwillende menselijke, criminele, terroristische of statelijke actoren nog toereikend wanneer AI-agents bij de uitvoering van een legitieme onderzoeks- of testopdracht zelfstandig schadelijke of ongeautoriseerde handelingswijzen kiezen die door hun ontwikkelaar niet waren beoogd? Zo ja, waarom? Zo nee, welke gevolgen heeft een gewijzigde beoordeling voor het nationale dreigingsbeeld en voor de taken en bevoegdheden van AIVD, MIVD, NCTV, NCSC en andere betrokken instanties?

Vraag 3
Acht u het juridisch en beleidsmatig nog toereikend om bij AI-dreigingen primair te zoeken naar de persoon, organisatie of staat die het systeem inzet, of moet daarnaast rekening worden gehouden met situaties waarin het onmiddellijke risico wordt veroorzaakt door autonoom systeemgedrag dat door de ontwikkelaar niet is beoogd en mogelijk niet volledig wordt beheerst?

Vraag 4
Kunt u, zonder mededelingen te doen over concrete operationele onderzoeken, aangeven of de incidenten rond OpenAI en Hugging Face inmiddels zijn betrokken bij analyses van AIVD, MIVD, NCTV en NCSC over toekomstige cyber- en nationale-veiligheidsdreigingen? Kunt u indien een nieuwe beoordeling nog niet heeft plaatsgevonden, aangeven of daartoe aanleiding bestaat nu OpenAI het incident publiek heeft bevestigd en daarover een technisch onderzoek heeft gepubliceerd?

II. AIVD, MIVD en de nieuwe Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten

Vraag 5
Artikel 2, tweede lid, onder a, van het conceptvoorstel[9] voor de Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten omschrijft de veiligheidstaak van de AIVD als het verrichten van onderzoek naar organisaties en personen die door de doelen die zij nastreven of door hun activiteiten aanleiding geven tot het ernstige vermoeden dat zij een gevaar vormen voor het voortbestaan van de democratische rechtsorde, de veiligheid of andere gewichtige belangen van de staat. Is deze actorgerichte taakomschrijving voldoende voor een situatie waarin de concrete bedreiging rechtstreeks voortvloeit uit autonoom gedrag van een AI-systeem

[9] Overheid.nl | Consultatie Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten

en niet kan worden vastgesteld dat een persoon of organisatie juist dat schadelijke gedrag heeft beoogd of aangestuurd?

Vraag 6
Indien u meent dat ook een dergelijke autonome AI-dreiging binnen de voorgestelde taakomschrijving van de AIVD kan worden onderzocht, kan u dan precies uiteenzetten wie of wat in zo'n geval onderwerp van het inlichtingenonderzoek is? Is dat de AI-onderneming, zijn dat bestuurders of ontwikkelaars, is dat het technische systeem of de infrastructuur waarop het draait, of een combinatie daarvan? Welke juridische betekenis heeft daarbij het ontbreken van een oogmerk bij de betrokken onderneming om schade aan de Nederlandse nationale veiligheid toe te brengen?

Vraag 7
In de memorie van toelichting bij het conceptvoorstel wordt in hoofdstuk 2, p.8 opgemerkt dat technologische ontwikkelingen, waaronder kunstmatige intelligentie, de nationale veiligheidsomgeving in toenemende mate beïnvloeden. Vervolgens wordt vastgesteld dat de onderzoeken van AIVD en MIVD die zich richten op de veiligheidstaak van oudsher meer een focus hadden op „personen en organisaties als actoren". Welke concrete consequenties heeft u uit deze constatering getrokken voor de wettelijke taakomschrijving van beide diensten? Waarom is de formulering waarin personen en organisaties centraal staan niettemin gehandhaafd en acht u deze formulering voldoende voor dreigingen die rechtstreeks voortvloeien uit autonoom gedrag van technologische systemen?

Vraag 8
Artikel 7, vijfde lid, onder a, van het conceptvoorstel maakt het, in samenhang met de definitie van inlichtingenonderzoek, mogelijk in aanvulling op de geïntegreerde aanwijzing een inlichtingenonderzoek te starten bij een onvoorziene of onmiddellijke dreiging tegen de nationale veiligheid, indien een aanpassing van de geïntegreerde aanwijzing redelijkerwijs niet kan worden gevergd. Kan deze bepaling worden toegepast indien bijvoorbeeld een autonoom AI-systeem onverwacht Nederlandse vitale infrastructuur begint aan te vallen? Of blijft voor een dergelijk spoedonderzoek vereist dat het onderzoek materieel kan worden gebracht onder een van de taken van artikel 2, tweede lid, onder a of b, dan wel artikel 3, tweede lid, onder a of b?

Vraag 9
Artikel 3, tweede lid, onder a, onder 3°, van het conceptvoorstel geeft de MIVD onder meer tot taak onderzoek te verrichten dat nodig is voor het treffen van maatregelen ter voorkoming van activiteiten die ten doel hebben de veiligheid of paraatheid van de krijgsmacht te schaden. Is deze formulering voldoende wanneer autonome AI-agents militaire netwerken binnendringen, militaire informatie verwerven of de inzetbaarheid van militaire systemen aantasten zonder dat op dat moment kan worden vastgesteld dat een menselijke of statelijke actor het doel heeft de veiligheid of paraatheid van de krijgsmacht te schaden? Welke wettelijke grondslag biedt het wetsvoorstel in zo'n situatie? (

Inleiding vraag 10a, 10b en 10c
De artikelen 55 tot en met 58 van het conceptvoorstel bevatten een bijzonder regime voor onderzoek naar organisaties met offensieve doelen en activiteiten. Artikel 55 ziet onder meer op een organisatie die toebehoort aan een land dat doelen nastreeft of activiteiten verricht die tegen Nederland of Nederlandse belangen zijn gericht, of waarvan aannemelijk is dat zij in opdracht van dat land handelt. Artikel 56 bevat een regeling voor bepaalde militaire organisaties, artikel 57 voorziet in

rechtmatigheidstoetsing van de aanwijzing en artikel 58 regelt de uitoefening van bijzondere bevoegdheden jegens aangewezen organisaties.

Vraag 10a
Hoe functioneert dit regime wanneer de concrete cyberdreiging wordt veroorzaakt door autonome AI-agents waarvan het handelen niet kan worden toegeschreven aan een organisatie, militaire organisatie of achterliggende staat als bedoeld in de artikelen 55 of 56?

Vraag 10b
Kan een autonoom AI-systeem zelf op enigerlei wijze binnen deze bepalingen worden gebracht, of vereist toepassing steeds een aanwijsbare organisatie?

Vraag 10c
Acht u deze wettelijke systematiek voldoende toekomstbestendig voor dreigingen waarbij de onmiddellijke gevaarzetting voortvloeit uit autonoom systeemgedrag?

Inleiding vraag 11a, 11b en 11c
Artikel 2, tweede lid, onder c, van het conceptvoorstel geeft de AIVD tot taak maatregelen te bevorderen ter bescherming van het voortbestaan van de democratische rechtsorde en van de veiligheid en andere gewichtige belangen van de staat. Hieronder vallen volgens onderdeel c, onder 2°, in ieder geval maatregelen ter beveiliging van onderdelen van de overheidsdienst en andere entiteiten die naar het oordeel van de verantwoordelijke ministers van vitaal belang zijn voor de instandhouding van het maatschappelijk leven.

Vraag 11a
Welke mogelijkheden biedt deze taak om banken, energievoorziening, gezondheidszorg, digitale infrastructuur en andere vitale processen preventief tegen autonome AI-aanvallen te beschermen?

Vraag 11b
Kan de AIVD op grond van deze bepaling ook maatregelen bevorderen wanneer nog geen vijandige menselijke, criminele, terroristische of statelijke actor is geïdentificeerd maar wel aanwijzingen bestaan voor een ernstige dreiging vanuit een autonoom AI-systeem?

Vraag 11c
Betreft deze bepaling uitsluitend het bevorderen van weerbaarheid bij mogelijke doelwitten, of kan zij ook een grondslag bieden voor maatregelen die zijn gericht op de onderneming die het betrokken AI-systeem ontwikkelt of exploiteert?

Vraag 12
Is bij de voorbereiding van het conceptwetsvoorstel specifiek onderzocht of de ontwikkeling van agentic AI en autonome cybercapaciteiten, waaronder de in 2026 bekend geworden ontwikkelingen, aanleiding geeft de taakomschrijvingen, onderzoeksbevoegdheden of spoedprocedures aan te passen? Zo nee, bent u bereid dit alsnog te doen voordat het wetsvoorstel na afloop van de consultatie aan de Afdeling advisering van de Raad van State wordt voorgelegd?

III. Preventief ingrijpen bij ondernemingen die zeer geavanceerde AI ontwikkelen

Vraag 13
Artikel 2, achtste lid, van de Europese AI-verordening bepaalt dat de verordening niet van toepassing is op onderzoeks-, test- of ontwikkelingsactiviteiten met

betrekking tot AI-systemen of AI-modellen vóór het moment waarop deze op de markt worden gebracht of in gebruik worden gesteld, waarbij testen onder reële omstandigheden van deze uitzondering is uitgesloten. Het bij het OpenAI/Hugging Face-incident betrokken model was juist een uitsluitend intern gebruikt onderzoeksmodel en de incidenten deden zich voor tijdens interne cybersecurity-evaluaties. Acht u het mogelijk dat hierdoor voor dit soort situaties een toezichtshiaat bestaat, doordat ernstige veiligheidsrisico's zich al tijdens interne onderzoeks-, ontwikkelings- en trainingsactiviteiten kunnen manifesteren voordat het normale marktgerichte regime van de verordening volledig toepasselijk wordt? Kunt u daarbij tevens ingaan op de betekenis van de laatste zin van artikel 2, achtste lid, volgens welke deze activiteiten wel overeenkomstig het overige toepasselijke Unierecht moeten worden verricht?

Inleiding vraag 14a en 14b
Volgens de NRC-berichtgeving van 11 september 2026 wist OpenAI aanvankelijk zelf niet dat zijn agents verantwoordelijk waren voor de inbraak bij Hugging Face.

Vraag 14a
Welke Nederlandse of Europese publieke instantie zou in een vergelijkbare situatie bevoegd zijn onafhankelijk onderzoek af te dwingen en de verdere experimenten zo nodig onmiddellijk te doen stilleggen?

Vraag 14b
Is een publieke autoriteit tijdens deze interne onderzoeks- en ontwikkelingsfase afhankelijk van vrijwillige medewerking van de betrokken AI-onderneming?

Vraag 15
Welke bevoegdheden bestaan thans om een AI-onderneming die medewerking weigert te verplichten een onafhankelijke publieke toezichthouder toegang te geven tot een nog niet op de markt gebracht model, de relevante testomgeving, veiligheidslogs, incidentgegevens en andere voor een veiligheidsbeoordeling noodzakelijke technische gegevens? Kan, indien strikt noodzakelijk en onder passende waarborgen voor bedrijfsgeheimen en intellectuele eigendom, ook toegang tot modelgewichten worden afgedwongen? Kunt u daarbij betrekken dat NRC op 13 september 2026 heeft bericht dat Anthropic in een ander geval toegang van het Britse AI Security Institute tot een nieuw model voor veiligheidstests had geweigerd?

Vraag 16
In de OpenAI-casus werden, blijkens het verslag OpenAI zelf en zoals te lezen in NRC van 11 september 2026, de experimenten uiteindelijk door de onderneming zelf stilgelegd nadat verdachte activiteiten waren ontdekt. Bestaat in Nederland of op Europees niveau een publieke autoriteit die in een vergelijkbaar geval hetzelfde juridisch kan afdwingen indien een onderneming niet zelf tot stillegging overgaat? Zo ja, welke autoriteit is dat, op welke wettelijke grondslag kan zij optreden en geldt deze bevoegdheid ook voor een nog niet op de markt gebracht intern onderzoeksmodel? Zo nee, acht u een dergelijke bevoegdheid noodzakelijk?

Vraag 17
Acht u het wenselijk dat voor de zwaarste categorie zeer geavanceerde AI-systemen een wettelijke verplichting wordt ingevoerd tot onafhankelijke veiligheidsevaluaties, waaronder adversarial testing, onderzoek naar autonome cybercapaciteiten, het vermogen veiligheidsmaatregelen te omzeilen, ongeautoriseerd internettoegang te verkrijgen en pogingen menselijke controle te ontwijken, voordat verdere training boven een nader vast te stellen risicodrempel mag plaatsvinden? Artikel 55 van de Europese AI-verordening bevat voor general-purpose AI-modellen

met systeemrisico reeds verplichtingen op het gebied van modelevaluatie, adversarial testing, beoordeling en beperking van systeemrisico's en cybersecurity. In hoeverre bieden deze verplichtingen volgens u voldoende waarborgen tijdens de ontwikkelings- en trainingsfase, en in hoeverre bestaat ruimte of noodzaak voor aanvullende nationale of Europese regels?

Vraag 18
Acht u daarnaast een wettelijke noodbevoegdheid wenselijk waarmee een bevoegde publieke toezichthouder bij ernstige en onderbouwde aanwijzingen voor acuut gevaar tijdelijk kan bevelen de training of verdere ontwikkeling van een model te staken, internet- of netwerktoegang van het model te beperken of te beëindigen, relevante gegevens, logs en andere onderzoeksgegevens veilig te stellen en verdere verspreiding of overdracht van modelgewichten tijdelijk te blokkeren? Aan welke wettelijke drempel, maximale duur, rechterlijke toetsing, periodieke herbeoordeling en bescherming van bedrijfsgeheimen zou een dergelijke bevoegdheid moeten worden gebonden?

Vraag 19
Acht u voor de hoogste risicocategorie tevens een voorafgaande meld- of vergunningplicht denkbaar, waarbij niet uitsluitend wordt gekeken naar de hoeveelheid gebruikte rekenkracht, maar tevens naar aantoonbare capaciteiten van het systeem, zoals autonomie, cybercapaciteiten, zelfreplicatie, het verkrijgen van nieuwe toegangsrechten, het omzeilen van toezicht en het vermogen zelfstandig andere agents aan te sturen?

Vraag 20
Artikel 9, eerste lid, van Verordening (EU) 2021/821 betreffende producten voor tweeërlei gebruik bepaalt dat een lidstaat om redenen van openbare veiligheid, waaronder het voorkomen van terroristische daden, de uitvoer van niet in bijlage I opgenomen producten voor tweeërlei gebruik kan verbieden of aan een vergunningplicht kan onderwerpen. Artikel 2, vijfde lid, van de Wet strategische diensten geeft de minister, onder verwijzing naar artikel 9, eerste lid, van deze verordening, de bevoegdheid om bij ministeriële regeling om redenen van openbare veiligheid of uit mensenrechtenoverwegingen een verbod of vergunningplicht in te stellen voor de overdracht van programmatuur of technologie als bedoeld in artikel 2, tweede lid, onderdeel d, van de verordening. Kan deze wettelijke combinatie naar uw oordeel worden gebruikt om de uitvoer of elektronische overdracht van nauwkeurig omschreven categorieën zeer geavanceerde AI-technologie aan een vergunningplicht of verbod te onderwerpen? Kunt u daarbij in het bijzonder ingaan op de vraag of modelgewichten, bepaalde trainingssoftware, gespecialiseerde software voor autonome AI-agents en andere voor de reproduceerbaarheid of verdere ontwikkeling van een zeer geavanceerd model essentiële digitale bestanden onder de begrippen programmatuur, technologie of product voor tweeërlei gebruik van deze regelgeving kunnen vallen? Zo nee, welke nationale wetswijziging of Europese aanpassing zou daarvoor noodzakelijk zijn?

Inleiding vraag 21a en 21b
De Regeling geavanceerde productieapparatuur voor halfgeleiders bevat een concreet voorbeeld van het gebruik van een dergelijke exportcontrolesystematiek. Artikel 1 van deze regeling verstaat onder geavanceerde productieapparatuur voor halfgeleiders niet alleen fysieke productieapparatuur, maar tevens de in de bijlage technisch omschreven programmatuur en technologie. Artikel 2 verbiedt uitvoer daarvan uit Nederland zonder vergunning van de minister. Artikel 3 bepaalt dat een vergunningaanvraag onder meer de bestemming en eindbestemming en de ontvanger en eindgebruiker moet vermelden. De bijlage bepaalt technisch welke apparatuur,

programmatuur en technologie onder de regeling vallen. Artikel 5a bepaalt dat de regeling mede berust op artikel 2, vijfde lid, van de Wet strategische diensten.

Vraag 21a
Acht u een vergelijkbare systematiek denkbaar voor bepaalde zeer geavanceerde AI-technologie? Zou de analogie daarin kunnen bestaan dat, zoals bij geavanceerde halfgeleidertechnologie een technisch omschreven categorie apparatuur, programmatuur en technologie niet zonder voorafgaande vergunning uit Nederland mag worden uitgevoerd, voor zeer geavanceerde AI een technisch en risicogebaseerd omschreven categorie modelgewichten, programmatuur of andere overdraagbare technologie wordt aangewezen waarvan de uitvoer of elektronische overdracht slechts na een veiligheidstoets is toegestaan? Zou bij die toets, naast bestemming, eindgebruiker en eindgebruik, specifiek kunnen worden gekeken naar het risico dat de technologie wordt gebruikt voor autonome cybercapaciteiten, het omzeilen van veiligheidsmaatregelen, zelfreplicatie of andere ernstige systeemrisico's?

Vraag 21b
Kunt u tevens aangeven waar de analogie met de halfgeleiderregeling ophoudt? Is het juist dat een exportcontroleregeling in beginsel de uitvoer of overdracht naar het buitenland reguleert en daarmee op zichzelf geen algemene grondslag biedt om de ontwikkeling, training of het gebruik van hetzelfde AI-model binnen Nederland stil te leggen? Welke aanvullende wettelijke grondslag zou daarvoor nodig zijn?

Vraag 22
Het kabinet heeft aangekondigd dat kunstmatige intelligentie per 1 januari 2027 wordt toegevoegd aan de sensitieve technologieën waarop de Wet veiligheidstoets investeringen, fusies en overnames van toepassing kan zijn. Is het juist dat deze wet mogelijkheden biedt om risico's voor de nationale veiligheid bij investeringen, fusies, overnames en verwerving van zeggenschap te beoordelen en zo nodig voorwaarden te stellen of een investering te verbieden, maar geen algemene bevoegdheid geeft om de ontwikkeling of training van een gevaarlijk AI-model door de onderneming zelf stil te leggen? Welke aanvullende wettelijke instrumenten staan in zo'n geval ter beschikking?

Vraag 23
Kunt u uiteenzetten of de Wet beschikbaarheid goederen in enig reëel AI-veiligheidsscenario een grondslag kan bieden voor overheidsingrijpen bij een AI-onderneming? Waar liggen de grenzen van deze wet wanneer het probleem niet bestaat uit het beschikbaar houden van essentiële goederen of productiecapaciteit, maar juist uit het veiligheidsrisico dat de ontwikkelde technologie zelf veroorzaakt? Kunt u daarbij betrekken dat de Wet beschikbaarheid goederen in 2025 bij Nexperia werd ingezet vanwege acute risico's voor het behoud en de continuïteit van cruciale technologische kennis en capaciteiten, terwijl het reguliere productieproces uitdrukkelijk kon doorgaan?[10]

IV. Mogelijke lacunes, wetswijziging en Europese actie

Vraag 24
Bent u bereid vóór indiening van het voorstel voor de Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten bij de Staten-Generaal een integrale juridische analyse op te stellen van de gevolgen van autonome AI-agents voor het Nederlandse stelsel van nationale veiligheid? Kan daarin afzonderlijk worden ingegaan op AI als hulpmiddel van staten, criminelen en terroristische

[10] Wet beschikbaarheid goederen ingezet door Minister van Economische Zaken | Rijksoverheid.nl

organisaties; autonoom AI-gedrag dat niet rechtstreeks door een menselijke actor wordt aangestuurd; de taken en bevoegdheden van AIVD en MIVD; de bescherming van vitale infrastructuur; de bevoegdheden van cyber- en AI-toezichthouders; de mogelijkheden onafhankelijk onderzoek bij AI-ondernemingen af te dwingen; en de mogelijkheden preventief in te grijpen bij ondernemingen die zeer geavanceerde AI ontwikkelen?

Vraag 25
Bent u bereid bij de advisering over het voorstel voor de Wet bescherming nationale veiligheid door de inlichtingen- en veiligheidsdiensten de Afdeling advisering van de Raad van State uitdrukkelijk aandacht te laten besteden aan de vraag of de voorgestelde taakomschrijvingen en bevoegdheden van AIVD en MIVD voldoende zijn toegesneden op nationale-veiligheidsdreigingen die voortvloeien uit autonoom opererende AI-systemen? Kan daarbij tevens worden gevraagd of het in het wetsvoorstel gehanteerde onderscheid tussen personen, organisaties, staten en technologische systemen voldoende toekomstbestendig is en of aanpassing of verduidelijking van de taakgrondslagen nodig is?

Vraag 26
Bent u tevens bereid de Afdeling advisering van de Raad van State om voorlichting te vragen over de ruimte die het Unierecht Nederland laat om de ontwikkeling en training van zeer risicovolle AI-systemen vóór marktintroductie nationaal aan een meldplicht, vergunningplicht, onafhankelijke veiligheidstoets of tijdelijke stilleggingsbevoegdheid te onderwerpen, in het bijzonder gelet op artikel 2, achtste lid, van de Europese AI-verordening en de harmoniserende werking van die verordening?

Vraag 27
Bent u bereid, indien voor de beantwoording van deze vragen nadere juridische of technisch-wetenschappelijke duidelijkheid nodig is, tevens advies van de landsadvocaat of andere onafhankelijke juridische en technisch-wetenschappelijke deskundigen in te winnen en de conclusies daarvan, voor zover mogelijk, aan de Staten-Generaal beschikbaar te stellen?

Vraag 28
Acht u aanpassing van de Europese AI-verordening wenselijk indien blijkt dat zeer geavanceerde AI-modellen reeds tijdens de interne onderzoeks-, ontwikkelings- en trainingsfase ernstige systeemrisico's kunnen veroorzaken? Bent u bereid in Europees verband aan de orde te stellen of de uitzondering van artikel 2, achtste lid, voor onderzoek, testen en ontwikkeling vóór marktintroductie voor modellen boven een nader vast te stellen risicodrempel moet worden beperkt of van aanvullende veiligheidsvoorwaarden moet worden voorzien?

Vraag 29
Acht u het, mede gelet op het grensoverschrijdende karakter van geavanceerde AI, noodzakelijk dat de Europese Unie afspraken maakt met onder meer de Verenigde Staten en het Verenigd Koninkrijk over verplichte melding van ernstige veiligheidsincidenten, onafhankelijke toegang voor veiligheidsonderzoek en procedures voor gecoördineerd ingrijpen wanneer een zeer geavanceerd model een onmiddellijk en ernstig internationaal veiligheidsrisico veroorzaakt?

Vraag 30
Kunt u aangeven of u voor de zwaarste categorie autonome AI een reguleringsmodel denkbaar acht waarbij het uitgangspunt niet uitsluitend is dat de overheid achteraf kan optreden, maar waarbij bepaalde ontwikkelingsactiviteiten vooraf worden

onderworpen aan melding, onafhankelijke veiligheidstoetsing of vergunningverlening? Zo nee, hoe kan naar haar oordeel met het bestaande en voorgenomen instrumentarium een vergelijkbaar preventief veiligheidsniveau worden bereikt?

V. Buitenlandse AI-ontwikkeling, Nederlandse aanknopingspunten en sanctiebevoegdheden

Inleiding vraag 31a, 31b, 31c en 31d
De Europese AI-verordening (EU) 2024/1689 bevat reeds - ongewijzigd gebleven - regels die ook ondernemingen buiten de Europese Unie onder haar bereik brengen. Artikel 2, eerste lid, onder a, bepaalt dat de verordening van toepassing is op aanbieders die AI-systemen in de Unie in de handel brengen of in gebruik stellen en op aanbieders die general-purpose AI-modellen in de Unie in de handel brengen, ongeacht of die aanbieders in de Unie of in een derde land zijn gevestigd. Artikel 2, eerste lid, onder c, brengt daarnaast buiten de Unie gevestigde aanbieders en gebruiksverantwoordelijken van AI-systemen onder het toepassingsgebied wanneer de door het AI-systeem geproduceerde output in de Unie wordt gebruikt. Artikel 53 verplicht aanbieders van general-purpose AI-modellen onder meer technische documentatie over het model, het trainings- en testproces en de evaluatieresultaten op te stellen en beschikbaar te houden en bepaalde informatie aan downstreamaanbieders beschikbaar te stellen. Artikel 55 stelt voor general-purpose AI-modellen met systeemrisico aanvullende verplichtingen inzake modelevaluatie, adversarial testing, beoordeling en beperking van systeemrisico's, ernstige incidenten en cybersecurity. Artikel 54 verplicht aanbieders uit derde landen bovendien, behoudens de daarin genoemde uitzondering, vóór het op de Uniemarkt brengen van een general-purpose AI-model een in de Unie gevestigde gemachtigde vertegenwoordiger aan te wijzen.

Vraag 31a
Kunt u tegen deze achtergrond uiteenzetten welke veiligheidsverplichtingen de Europese AI-verordening thans reeds kan opleggen aan een onderneming die een zeer geavanceerd model geheel buiten de Europese Unie ontwikkelt en traint, maar dat model vervolgens op de Uniemarkt brengt?

Vraag 31b
In hoeverre kunnen de verplichtingen uit de verordening mede betrekking hebben op risico's die hun oorsprong vinden in de voorafgaande buitenlandse ontwikkeling en training?

Vraag 31c
Welke relevante veiligheidsrisico's van deze buitenlandse onderzoeks- en trainingsfase vallen volgens u niet of niet volledig onder de bestaande AI-verordening, mede gelet op artikel 2, achtste lid?

Vraag 31d
Welke aanvullende bevoegdheden kan Nederland in zoverre zelfstandig vaststellen, in het bijzonder indien de buitenlandse ontwikkeling of training een bedreiging voor de Nederlandse nationale veiligheid kan opleveren?

Vraag 32
Acht u het juridisch mogelijk om als voorwaarde voor toegang tot de Nederlandse of Europese markt te verlangen dat de aanbieder van een zeer geavanceerd AI-model aantoont dat ook de ontwikkeling, training en veiligheidstests buiten Nederland en buiten de Europese Unie aan bepaalde essentiële veiligheidsnormen hebben voldaan? Welke ruimte laat de Europese AI-verordening voor aanvullende nationale voorwaarden

en welke onderdelen zouden vanwege de harmoniserende werking van de verordening of het grensoverschrijdende karakter uitsluitend of bij voorkeur op het niveau van de Europese Unie moeten worden geregeld?

Inleiding vraag 33a, 33b en 33c
Artikel 54, eerste lid, van de Europese AI-verordening verplicht een in een derde land gevestigde aanbieder van een general-purpose AI-model, behoudens de uitzondering van het zesde lid, vóór het op de Uniemarkt brengen van het model schriftelijk een in de Unie gevestigde gemachtigde vertegenwoordiger aan te wijzen. Op grond van het derde lid moet het mandaat deze vertegenwoordiger onder meer in staat stellen te controleren of de voorgeschreven technische documentatie is opgesteld en of aan de verplichtingen van artikel 53 en, indien toepasselijk, artikel 55 is voldaan, documentatie beschikbaar te houden, informatie aan het AI-bureau te verstrekken en met het AI-bureau en bevoegde autoriteiten samen te werken.

Vraag 33a
Kunt u uiteenzetten welke betekenis deze constructie kan hebben voor effectief toezicht op zeer geavanceerde AI-modellen van buiten de Europese Unie gevestigde ondernemingen? Kan de gemachtigde vertegenwoordiger een juridisch aanknopingspunt vormen voor veiligheidsverplichtingen die mede betrekking hebben op voorafgaande buitenlandse ontwikkeling, training en veiligheidstests?

Vraag 33b
Kunt u daarbij met name het bepaalde in artikel 54, vierde lid, betrekken? Dit lid bepaalt dat het mandaat de gemachtigde vertegenwoordiger ertoe moet machtigen om naast of in plaats van de buitenlandse aanbieder door het AI-bureau of de bevoegde autoriteiten te worden aangesproken over alle vraagstukken die verband houden met de naleving van de AI-verordening. Welke juridische en praktische mogelijkheden biedt deze bepaling reeds om een buiten de Europese Unie gevestigde AI-ontwikkelaar via zijn Europese vertegenwoordiger daadwerkelijk tot naleving te bewegen? Hoever strekt deze verantwoordelijkheid van de vertegenwoordiger en waar liggen de grenzen daarvan?

Vraag 33c
Acht u het mogelijk, indien voor zeer geavanceerde AI aanvullende veiligheidsverplichtingen worden ingevoerd, op deze bestaande constructie voort te bouwen door te bepalen dat de gemachtigde vertegenwoordiger tevens aanspreekpunt is voor verplichtingen inzake ernstige veiligheidsincidenten, onafhankelijke veiligheidsevaluaties, toegang tot relevante documentatie en modellen en de naleving van eventuele bevelen tot beperking of tijdelijke stillegging? Zou daarvoor wijziging van de Europese AI-verordening nodig zijn of bestaat voor bepaalde nationale-veiligheidsmaatregelen ruimte voor aanvullende Nederlandse wetgeving?

Vraag 34
Welke ruimte biedt het nationale en Europese recht om bij een zeer ernstig veiligheidsrisico niet alleen het betrokken AI-model van de Nederlandse of Europese markt te weren, maar tevens nauw met dat model verbonden diensten van de betrokken onderneming geheel of gedeeltelijk te beperken zolang zij een bindend veiligheidsbevel niet naleeft? Welke grenzen stellen het Unierecht, het evenredigheidsbeginsel, het eigendomsrecht en de vrijheid van ondernemerschap aan dergelijke maatregelen?

Vraag 35

Acht u het juridisch mogelijk om in Nederland gevestigde datacenter- en cloudexploitanten te verplichten geen rekenkracht, opslagcapaciteit, netwerkverbindingen of andere technische voorzieningen beschikbaar te stellen voor de training, verdere ontwikkeling of exploitatie van een zeer geavanceerd AI-model waarvan een bevoegde autoriteit heeft vastgesteld dat het een ernstig en onvoldoende beheerst gevaar voor de nationale veiligheid vormt? Kan daarbij onderscheid worden gemaakt tussen een bevel betreffende een concreet model en een bredere beperking van dienstverlening aan de betrokken onderneming?

Vraag 36
Kan een bevoegde Nederlandse autoriteit onder de huidige wetgeving een in Nederland gevestigd datacenter of cloudbedrijf bevelen de dienstverlening ten aanzien van een bepaalde buitenlandse AI-ontwikkelaar of een bepaald AI-model te beëindigen wanneer de gevaarlijke ontwikkeling of training zelf geheel buiten Nederland plaatsvindt? Zo ja, op welke wettelijke grondslag, bij welke gevaarzetting en onder welke procedurele voorwaarden? Zo nee, acht u het juridisch mogelijk en wenselijk daarvoor een wettelijke bevoegdheid te creëren?

Inleiding vraag 37a en 37b
Een aantal grote technologieconcerns beschikt over Nederlandse of Europese dochtermaatschappijen, vestigingen of andere rechtspersonen, terwijl de ontwikkeling van hun meest geavanceerde modellen elders binnen het concern kan plaatsvinden.

Vraag 37a
In hoeverre kan naar huidig recht een in Nederland of elders in de Europese Unie gevestigde dochtermaatschappij, vestiging of gemachtigde vertegenwoordiger worden aangesproken voor veiligheidsverplichtingen met betrekking tot activiteiten van een buitenlandse moeder- of zustermaatschappij?

Vraag 37b
Acht u een wettelijke regeling denkbaar waarbij een concern dat zeer geavanceerde AI op de Nederlandse of Europese markt aanbiedt verplicht wordt een in de Europese Unie gevestigde rechtspersoon aan te wijzen die voor bepaalde veiligheidsverplichtingen daadwerkelijk juridisch kan worden aangesproken? Kan daarbij worden voortgebouwd op de constructie van artikel 54 van de Europese AI-verordening?

Vraag 38
Acht u een getrapt sanctie- en handhavingsstelsel juridisch mogelijk waarbij overtreding van bindende veiligheidsvoorschriften achtereenvolgens of afhankelijk van de ernst kan leiden tot een informatie- of medewerkingsbevel, een bestuurlijke boete, tijdelijke beperking van het aanbieden van het betrokken AI-model, een verbod op het gebruik van Nederlandse cloud- of datacenterinfrastructuur en, in het uiterste geval, tijdelijke beperking van nader omschreven andere activiteiten van het betrokken concern in Nederland? Welke van deze bevoegdheden kunnen door Nederland zelfstandig worden vastgesteld en voor welke is regelgeving op het niveau van de Europese Unie noodzakelijk?

Vraag 39
Welke mogelijkheden biedt het recht van de Europese Unie om financiële of andere economische beperkende maatregelen te treffen tegen een buitenlandse onderneming die ondanks een bindend besluit doorgaat met activiteiten met zeer geavanceerde AI waarvan is vastgesteld dat zij een ernstige bedreiging voor de nationale of Europese veiligheid veroorzaken? Kunt u daarbij in het bijzonder ingaan op artikel

215, tweede lid, van het Verdrag betreffende de werking van de Europese Unie, dat na een voorafgaand besluit in het kader van het gemeenschappelijk buitenlands en veiligheidsbeleid beperkende maatregelen tegen natuurlijke of rechtspersonen, groepen en niet-statelijke entiteiten mogelijk maakt? In hoeverre zou een ernstig AI-veiligheidsrisico onder omstandigheden aanleiding kunnen zijn te onderzoeken of deze Europese route bruikbaar of aanpasbaar is?

Vraag 40
Is het naar uw oordeel juridisch mogelijk dat Nederland zelfstandig Nederlandse banken, beleggingsondernemingen en andere financiële dienstverleners verbiedt transacties te verrichten in aandelen, obligaties of andere financiële instrumenten van een buitenlandse onderneming die ernstige en bindende AI-veiligheidsvoorschriften overtreedt? Kunt u daarbij ingaan op artikel 63 VWEU, dat in beginsel beperkingen op het kapitaal- en betalingsverkeer tussen lidstaten en tussen lidstaten en derde landen verbiedt, en artikel 65 VWEU, dat onder voorwaarden maatregelen toelaat die onder meer worden gerechtvaardigd door openbare orde of openbare veiligheid, mits deze geen willekeurige discriminatie of verkapte beperking van het vrije kapitaalverkeer vormen?

Vraag 41
Is het juist dat de bestaande Europese bevoegdheden tot schorsing of verwijdering van financiële instrumenten uit de handel op een gereglementeerde markt in hoofdzaak zijn vormgegeven vanuit de regels van het handelsplatform, de bescherming van beleggers en de ordelijke werking van de financiële markt en niet als algemene nationale-veiligheidssanctie wegens activiteiten van de uitgevende onderneming op een ander terrein? Indien dit juist is, acht u aanvullende nationale of Europese wettelijke bevoegdheden denkbaar of noodzakelijk indien men effectenhandel als uiterste veiligheidsmaatregel zou willen kunnen beperken?

Vraag 42
Kunt u uiteenzetten of de Sanctiewet 1977 in haar huidige vorm een grondslag kan bieden voor financiële of andere economische maatregelen tegen een private buitenlandse AI-onderneming die door ontwikkeling, training of exploitatie van zeer geavanceerde AI-systemen een ernstige bedreiging voor de Nederlandse nationale veiligheid veroorzaakt, indien tegen die onderneming geen internationale of Europese sanctiemaatregel bestaat? Is het juist dat artikel 2 van de Sanctiewet 1977 de mogelijkheid tot het vaststellen van sanctieregels koppelt aan verdragen, besluiten of aanbevelingen van organen van volkenrechtelijke organisaties en internationale afspraken met betrekking tot onder meer de internationale vrede en veiligheid, de internationale rechtsorde of terrorismebestrijding, terwijl artikel 3 bepaalt op welke terreinen dergelijke regels kunnen worden gesteld, waaronder goederen-, diensten- en financieel verkeer?

Vraag 43
Heeft u bij de voorbereiding van het bij de Tweede Kamer aanhangige voorstel voor de Wet internationale sanctiemaatregelen[11] onderzocht of Nederland daarnaast behoefte heeft aan een wettelijke bevoegdheid om in uitzonderlijke gevallen eigenstandige nationale sanctiemaatregelen te treffen tegen rechtspersonen wier activiteiten een ernstige bedreiging voor de nationale veiligheid opleveren, zonder dat daaraan reeds een internationale of Europese sanctiemaatregel ten grondslag ligt? Zo ja, tot welke conclusie heeft dit onderzoek geleid? Zo nee, bent u bereid dit alsnog te onderzoeken? Kunt u daarbij betrekken dat de memorie van toelichting ook de voorgestelde Wet internationale sanctiemaatregelen omschrijft als een

[11] *Kamerstukken II,* 2025-2026, 36.898, 2 voorstel van wet Regels over de uitvoering van internationale sanctiemaatregelen (Wet internationale sanctiemaatregelen)

kaderwet voor maatregelen die Nederland in internationaal verband is overeengekomen?

Vraag 44
Acht u het juridisch mogelijk en wenselijk om de Sanctiewet 1977, het voorstel voor de Wet internationale sanctiemaatregelen dan wel andere nationale wetgeving uit te breiden met een uitdrukkelijke grondslag voor eigenstandige nationale veiligheidsmaatregelen tegen rechtspersonen, waaronder ondernemingen die zeer geavanceerde AI-systemen ontwikkelen of exploiteren en wier activiteiten een ernstige bedreiging voor de nationale veiligheid vormen? Welke beperkingen vloeien daarbij voort uit het recht van de Europese Unie, het internationale recht, de territoriale reikwijdte van de Nederlandse rechtsmacht, het vrije verkeer van kapitaal en diensten en de bevoegdheidsverdeling tussen Nederland en de Europese Unie?

Vraag 45
Welke soorten maatregelen zouden, indien voor een dergelijke wettelijke grondslag wordt gekozen, juridisch denkbaar zijn? Kunt u daarbij afzonderlijk ingaan op het bevriezen van in Nederland aanwezige vermogensbestanddelen; het verbieden of beperken van financiële dienstverlening en transacties; het verbieden van bepaalde investeringen; het beperken van dienstverlening door Nederlandse ondernemingen; het verbieden van het gebruik van Nederlandse cloud-, reken- of datacentercapaciteit; en maatregelen ten aanzien van Nederlandse vestigingen of dochtermaatschappijen van het betrokken concern? Kunt u daarbij aangeven welke van deze maatregelen Nederland zelfstandig zou kunnen treffen en welke vanwege het Unierecht slechts of effectiever op Europees niveau kunnen worden geregeld? De huidige Sanctiewet 1977 laat bij de uitvoering van internationale sanctiemaatregelen reeds regels toe ten aanzien van onder meer goederen-, diensten- en financieel verkeer.

Vraag 46
Acht u het juridisch denkbaar om onder een dergelijke regeling tevens beperkingen te stellen aan de aankoop, verkoop of bemiddeling in aandelen, obligaties of andere effecten van een betrokken onderneming? Kan daarbij duidelijk onderscheid worden gemaakt tussen een verbod voor Nederlandse financiële instellingen om bepaalde diensten te verlenen, een verbod voor Nederlandse personen of instellingen om bepaalde financiële instrumenten te verwerven en een daadwerkelijke schorsing of beëindiging van een beursnotering die zich bijvoorbeeld op een Amerikaanse beurs bevindt? Welke beperkingen volgen daarbij uit de artikelen 63 en 65 VWEU, de Europese financiële-marktwetgeving en het proportionaliteitsbeginsel?

Vraag 47
Welke materiële en procedurele waarborgen zouden volgens u noodzakelijk zijn voordat een onderneming aan dergelijke zware nationale-veiligheidsmaatregelen kan worden onderworpen? Kunt u daarbij ingaan op de vereiste ernst en waarschijnlijkheid van het gevaar, de kwaliteit van het technisch-wetenschappelijke bewijs, het causale verband tussen de activiteiten van de onderneming en de bedreiging voor Nederland, hoor en wederhoor, bescherming van bedrijfsgeheimen, maximale duur, periodieke herbeoordeling, onafhankelijke toetsing en effectieve rechterlijke rechtsbescherming?

Vraag 48
Indien u van oordeel bent dat een eigenstandig Nederlands sanctieregime juridisch niet mogelijk, niet doeltreffend of vanwege het grensoverschrijdende karakter van de betrokken ondernemingen niet wenselijk is, kunt u uiteenzetten welke van de hiervoor bedoelde economische en financiële maatregelen wel op het niveau van de

Europese Unie kunnen worden getroffen? Acht u het wenselijk te onderzoeken of op Europees niveau een specifiek regime van beperkende economische of financiële maatregelen nodig is voor ondernemingen die ondanks bindende veiligheidsvoorschriften doorgaan met de ontwikkeling of exploitatie van AI-systemen waarvan is vastgesteld dat zij een ernstige bedreiging voor de nationale of Europese veiligheid veroorzaken? Kan daarbij tevens de mogelijke betekenis en de beperkingen van artikel 215 VWEU worden betrokken?

Vraag 49
Bent u bereid de Afdeling advisering van de Raad van State om voorlichting te vragen over de vraag hoever Nederland zijn wetgeving territoriaal en extraterritoriaal kan laten reiken ten aanzien van zeer risicovolle AI-ontwikkeling door buitenlandse ondernemingen? Kan daarbij afzonderlijk worden onderzocht welke juridische ruimte bestaat voor voorwaarden voor Nederlandse markttoegang die mede betrekking hebben op buitenlandse ontwikkeling en training; verplichtingen voor Nederlandse of Europese concernonderdelen en gemachtigde vertegenwoordigers; verplichtingen voor in Nederland gevestigde datacenter- en cloudbedrijven; bestuurlijke en financiële maatregelen tegen buitenlandse moedermaatschappijen; beperkingen van financiële dienstverlening, investeringen en effectenhandel; en uitbreiding van de Sanctiewet 1977, de voorgestelde Wet internationale sanctiemaatregelen of andere nationale wetgeving met een grondslag voor eigenstandige nationale veiligheidsmaatregelen? Kan de Afdeling daarbij tevens worden gevraagd in te gaan op de verhouding tot de Europese AI-verordening, het vrije verkeer van kapitaal en diensten, het eigendomsrecht, de vrijheid van ondernemerschap, het internationale recht en de bevoegdheidsverdeling tussen Nederland en de Europese Unie?

Vraag 50
Bent u bereid te onderzoeken of de Europese AI-verordening of andere Europese regelgeving zo kan worden aangepast dat aanbieders van zeer geavanceerde AI-modellen die toegang tot de Uniemarkt wensen moeten aantonen dat hun relevante ontwikkeling, training en veiligheidstests wereldwijd aan nader vast te stellen minimumveiligheidsnormen hebben voldaan, ongeacht het land waarin deze activiteiten feitelijk plaatsvinden? Kan daarbij worden voortgebouwd op de reeds bestaande extraterritoriale aanknopingspunten van artikel 2 van de AI-verordening en op de in artikel 54 neergelegde constructie van de gemachtigde vertegenwoordiger?

Vraag 51
Acht u het wenselijk dat Nederland en de Europese Unie de regulering van zeer geavanceerde AI zodanig vormgeven dat een onderneming de werking daarvan niet kan ontwijken door de ontwikkeling en training van haar krachtigste modellen buiten Nederland of buiten de Europese Unie te laten plaatsvinden, terwijl zij tegelijkertijd via markttoegang, vestigingen, dochtermaatschappijen, gemachtigde vertegenwoordigers, datacenters, cloudinfrastructuur, digitale diensten, financiële dienstverlening of andere economische activiteiten substantieel met Nederland en de Europese Unie verbonden blijft? Zo ja, welke wettelijke instrumenten acht u daarvoor nodig en op welk niveau – nationaal, Europees of internationaal – dienen deze naar uw oordeel te worden vastgesteld?