Signaalbrief ‘Begin vandaag met de kennissprong voor AI-gestuurde cyberaanvallen’

Aan: Minister van Justitie en Veiligheid en Staatssecretaris Digitale Economie en Soevereiniteit
c.c.: Minister van Economische Zaken en Klimaat, Minister van Defensie, Minister van Buitenlandse Zaken, Staatssecretaris Koninkrijksrelaties en Slagvaardige Overheid

Excellenties,

Nederland is niet voorbereid op een nieuwe generatie cyberaanvallen waarbij kunstmatige intelligentie als wapen wordt ingezet. De Cyber Security Raad (hierna de raad) maakt zich hierover grote zorgen. Deze aanvallen zijn sneller, grootschaliger en moeilijker te detecteren dan traditionele cyberaanvallen en worden nu al autonoom ingezet door statelijke actoren. Onze verdediging loopt achter. In deze signaalbrief onderbouwt de raad deze zorgen en doet drie concrete aanbevelingen.

Samenvatting van het advies

De raad adviseert:

- AI-gestuurde cyberaanvallen als volwaardige strategische dreiging te erkennen en beleid, budgetten en operationele plannen daar zo snel mogelijk op aan te passen;
- De kennisachterstand van Nederland en de EU op dit gebied in de komende jaren substantieel in te lopen. Op andere terreinen is een zogenaamde *Grand Challenge* hiervoor een goed middel gebleken: een competitie gericht op het stimuleren van kennisopbouw, waarbij (multidisciplinaire) teams werken aan mogelijke *dual use*-oplossingen. De raad biedt aan om de organisatie hiervan, samen met bedrijfsleven en wetenschap, in Europees verband te onderzoeken.

Aanleiding

De gevolgen van de dreigingen van AI-gestuurde cyberaanvallen zijn urgent en de gevolgen voor de nationale en maatschappelijke veiligheid kunnen groot zijn. Hoewel geavanceerde cyberaanvallen met AI nog beperkt zijn, ontwikkelen veel landen in hoog tempo kennis en toepassingen. Het is dan ook de verwachting dat dergelijke aanvallen in de nabije toekomst een grote dreiging gaan vormen. Het inlopen van onze kennisachterstand op dit gebied is daarom dringend noodzakelijk. Deze kennisachterstand van Nederland neemt iedere dag toe en naarmate langer wordt gewacht met kennisopbouw en het ontwikkelen van toepassingen zal de kans op schade door dergelijke aanvallen toenemen en is er meer herstelcapaciteit nodig.

Deze zorgpunten gelden niet alleen voor Nederland, maar worden breed gedragen. Zo waarschuwen vooraanstaande onderzoekers in België, Duitsland en de Verenigde Staten voor de strategische implicaties van de kennisskloof voor westerse democratieën. Ook de Europese Commissie waarschuwt expliciet voor AI-gestuurde cyberwapens in het in oktober 2025 gepubliceerde rapport *‘Apply AI Strategy’*.

> *“Terrorist and organised criminal organisations are increasingly using AI-based technologies to accelerate, upscale and broaden the reach of their illicit activities. Cybercrime, sabotage and terrorism are blended into hybrid attacks, where AI is often exploited by malicious actors. We therefore need to ensure the swift delivery of AI-based solutions for internal security and cyber security.”*

Ontwikkelingen

De raad constateert vier ontwikkelingen over het gebruik van AI bij cyberaanvallen:

1. Cyberaanvallen met AI vinden in de praktijk al plaats

- AI-systemen kunnen snel, grootschalig en autonoom kwetsbaarheden identificeren, analyseren en exploiteren. Hoewel deze kwalitatief geavanceerde aanvallen nu nog beperkt in omvang zijn, hebben deze ontwikkelingen het potentieel om voor maatschappelijke ontwrichting te zorgen. Concrete voorbeelden illustreren de urgentie:
- Een AI-gestuurde ransomware-aanval op een ziekenhuisketen in de Verenigde Staten in januari 2025 leidde tot het betalen van 22 miljoen dollar aan losgeld. De inzet van AI zorgde ervoor dat de diefstal van patiëntengegevens onzichtbaar bleef voor zowel menselijke als geautomatiseerde beveiligingssystemen.
- Een Russische staatsgelieerde hackersgroep (APT28) zette in 2025 verfijnde, adaptieve malware in (LAMEHUG), waarbij AI eenvoudige tekstinstructies omzette naar gerichte cyberaanvallen op computers van slachtoffers.
- 2025 was ook het jaar waarin een aan China gelieerde hackersgroep een cyberspionage-aanval uitvoerde waarbij AI autonoom tientallen organisaties (technologiebedrijven, financiële instellingen en overheidsorganisaties) over de hele wereld aanviel.
- AI-technologie wordt steeds makkelijker in te zetten voor criminele activiteiten. Oplossingen voor cyberaanvallen vergelijkbaar met ChatGPT verschijnen aan de lopende band: FraudGPT, WormGPT, DarkGPT en vergelijkbare tools worden ingezet voor zowel grootschalige autonome aanvallen als kwalitatief geavanceerde aanvallen.
- AI-gestuurde scanning en malware verkorten de tijd tussen de ontdekking van een kwetsbaarheid en het misbruik ervan drastisch. Organisaties hebben daardoor minder tijd om kwetsbaarheden te verhelpen en zien zich vaker genoodzaakt tot overhaaste updates -zoals zichtbaar werd tijdens het Crowdstrike-incident in 2024, waarbij een geforceerde update wereldwijd tot grote verstoringen leidde.

2. Huidige specialistische kennis is onvoldoende beschikbaar

De expertise over geavanceerde AI-toepassingen voor cyberaanvallen is momenteel geconcentreerd in handen van kwaadwillende actoren - veelal staten met een offensief cyberprogramma- en inlichtingen- en veiligheidsdiensten van andere landen. Alleen wie begrijpt hoe een tegenstander opereert, kan zich daartegen effectief beschermen. Deze actoren delen deze kennis echter niet of nauwelijks, zelfs niet met bevriende naties of binnen samenwerkingsverbanden zoals de NAVO. Nederland ontbeert dan ook op dit moment de vereiste specialistische kennis om de aard, reikwijdte en methodieken van AI-gestuurde cyberaanvallen te doorgronden. Het zelf kunnen ontwikkelen van voldoende defensieve oplossingen en toepassingen ligt nog verder buiten bereik. Om niet alleen de kennisachterstand in te lopen, maar bij te blijven s het noodzakelijk om structureel te investeren in kennisontwikkeling en innovatie rondom dit thema. Landen buiten de EU zijn hier al volop mee bezig.

3. Defensieve slagkracht schiet tekort

De weerbaarheid tegen AI-gestuurde aanvallen vergt - naast kennis over de aard van dergelijke aanvallen - steeds vaker ook de inzet van geautomatiseerde systemen om aanvallen sneller te detecteren, misbruik te voorspellen en effectievere herstelmethoden te ontwikkelen. Dit zijn systemen waarin AI een sleutelrol speelt. Het Syzbot-project, door Google gelanceerd voor defensieve doeleinden, illustreert dit. Door middel van een genetisch algoritme (een oudere vorm van AI-technologie), wordt hierbij voortdurend gezocht naar kwetsbaarheden in gangbare besturingssystemen. Hoewel de AI-toepassing in Syzbot weinig geavanceerd is, vindt het systeem nu al meer bugs dan programmeurs kunnen verwerken: voortdurend wachten honderden tot duizenden softwarefouten op herstel.

Hoewel er rekenkracht en expertise in Nederland en de EU aanwezig zijn, onder meer in toekomstige AI-fabrieken, onderzoeksinstituten en bij universiteiten, ontbreekt een duidelijke strategie om deze ten dienste te stellen van onze digitale weerbaarheid. Het innovatieklimaat, de investeringen en de bedrijvigheid op dit gebied blijven achter, zeker vergeleken met andere toonaangevende landen.

Aanbevelingen

De raad heeft drie aanbevelingen om de achterstand op het gebied van cyberaanvallen met AI in te lopen en de dreiging het hoofd te bieden:

1. **Verwerk AI-gestuurde dreigingen in overheidsstrategieën en beleid.** Erken AI-gestuurde cyberaanvallen als volwaardige strategische dreiging en pas beleid, R&D-budgetten en operationele plannen daarop aan. Voer dit uit in publiek-private en wetenschappelijke samenwerking, waar mogelijk voortbordurend op bestaande coalities. Wacht niet op een volgende Nederlandse Cybersecuritystrategie (NLCS): begin nu.
2. **Investeer strategisch in het Nederlandse bedrijfsleven op het gebied van AI en cybersecurity.** Vergroot de defensieve slagkracht door innovatieve producten te ontwikkelen voor betere detectie van aanvallen met AI, verfijndere penetratietesten — ook voor ethische hackers — en gereedschappen voor actieve cyberverdediging. Benut hiervoor de investerings- en innovatie-instrumenten uit het coalitieakkoord Jetten-I.
3. **Verzoek om uitspreken van politieke steun voor een Europese *Grand Challenge* op het gebied van AI en cybersecurity.** Een Grand Challenge is een competitie die kennisopbouw en innovatie stimuleert door multidisciplinaire teams van onderzoekers en bedrijven — waaronder start-ups en scale-ups — te mobiliseren. Door de oplossingen als open-source aan te bieden, zijn ze in potentie commercieel schaalbaar binnen overheidsorganisaties én het bedrijfsleven. Binnen de EU is een challenge van deze omvang op het gebied van cybersecurity en AI nog niet eerder gelanceerd. Bedrijfsleven en wetenschap vertegenwoordigd in de raad, zien kansen om dit in Europees verband te organiseren en willen dit onderzoeken. Uw politieke steun kan bijdragen dit initiatief tot een succes te maken.

De raad roept het kabinet op om nog in 2026 concrete stappen te zetten op elk van de drie bovenstaande aanbevelingen. Uitstel vergroot de kwetsbaarheid van Nederland.

De raad is van harte bereid in gesprek te gaan over onze aanbevelingen.

Hoogachtend,

Marc Kuipers
Covoorzitter publieke sector

Sylvia van Es
Covoorzitter private sector