RAPPORT ANALYSE

Geavanceerde AI-modellen vergroten cyberrisico's voor ondernemingen

In het kort: Nieuwe, geavanceerde AI-modellen vergroten de kans op cyberaanvallen. Vooral kleine en middelgrote ondernemingen zijn kwetsbaar. Deze technologie maakt het makkelijker en sneller om zwakke plekken te vinden en te misbruiken, waardoor er minder tijd is om in te grijpen. Dit rapport laat zien wat dit betekent voor ondernemingen onder AFM-toezicht. De focus ligt op het versterken van de cyberweerbaarheid, onder meer door tijdig beveiligingsupdates door te voeren, goede monitoring, snelle incidentrespons en het op orde brengen van basismaatregelen.

JUNI | 2026

AFM

1. Aanleiding

Recente publicaties[1] en evaluaties[2] laten zien dat geavanceerde AI-modellen (General Purpose AI, zoals bedoeld in de EU AI-verordening) met toenemende cybercapaciteiten steeds beter in staat zijn om kwetsbaarheden te identificeren, aanvalsstappen te combineren en zelfstandig complexe cybertaken uit te voeren. Hierdoor wordt de tijd tussen het ontstaan of ontdekken van een kwetsbaarheid en mogelijk misbruik steeds korter[3].

Ontwikkelingen rond deze modellen illustreren de versnelling in het dreigingslandschap. Geavanceerde AI-modellen kunnen geautomatiseerd kwetsbaarheden opsporen en exploiteren en in gecontroleerde omgevingen geavanceerde aanvallen uitvoeren. Tegelijkertijd ontwikkelen ook andere, waaronder open-source, AI-modellen vergelijkbare capaciteiten, waardoor deze ontwikkeling breder relevant is. Hoewel deze modellen door ondernemingen en toezichthouders nog niet altijd volledig zelfstandig kunnen worden getest[4], laten onafhankelijke onderzoeken zien dat zij in gecontroleerde omgevingen steeds beter in staat zijn om complexe aanvalsstappen uit te voeren.

De AFM constateert dat dezelfde AI zowel bedreigend als defensief gebruikt wordt. Technologiebedrijven zetten diezelfde technologie bovendien in voor defensieve doeleinden, zoals het beschermen van vitale infrastructuur. Dit benadrukt dat AI zowel kansen biedt voor versterking van cyberweerbaarheid als risico's met zich meebrengt door mogelijk misbruik[5].

De mogelijkheden van AI-gedreven cyberaanvallen nemen snel toe[6]. Aanvallen worden complexer, kunnen langer autonoom doorgaan en zijn eenvoudiger schaalbaar[7]. Dit kan aanvallers op korte termijn een voorsprong geven[8]. Ook het Nationaal Cyber Security Centrum (NCSC-NL) waarschuwt[9] dat deze ontwikkeling cyberaanvallen aanzienlijk versnelt.

Tegen deze achtergrond neemt de druk op de cyberweerbaarheid van ondernemingen toe[10]. Van ondernemingen wordt verwacht dat zij hun basisbeveiliging op orde brengen, kwetsbaarheidsbeheer versterken en beveiligingsupdates sneller doorvoeren. Ook is het van belang dat zij hun monitoring en incidentrespons verbeteren en reactietijden verkorten. De AFM vraagt aandacht voor deze ontwikkeling bij alle ondernemingen en in het bijzonder bij middelgrote en kleinere partijen.

De AFM volgt deze ontwikkelingen nauwgezet en verwacht dat ondernemingen beoordelen welke aanvullende maatregelen binnen de eigen organisatie nodig zijn. Tegelijkertijd moedigt de AFM aan om te verkennen hoe AI kan worden ingezet om de eigen cyberweerbaarheid en die binnen de keten te versterken.

[1] Anthropic – Claude Mythos Preview / Red Team analyse
[2] Our evaluation of Claude Mythos Preview's cyber capabilities | AISI Work
[3] Anthropic – Project Glasswing
[4] Anthropic – Project Glasswing
[5] Generatieve AI: een transformatieve impact op cybersecurity | Rijksinspectie Digitale Infrastructuur (RDI)
[6] UK AI Security Institute – autonome AI-cybercapaciteiten
[7] Our evaluation of Claude Mythos Preview's cyber capabilities | AISI Work
[8] SANS Critical Advisory: BugBusters - AI Vulnerability Discovery Hype vs. Reality | SANS Institute
[9] Anthropic's frontiermodel Mythos vraagt om directe actie | NCSC
[10] CERT-EU - AI is changing the economics of vulnerability discovery. Defenders should adapt now

2. Wat vraagt dit van ondernemingen

Deze ontwikkeling vergroot de druk op ondernemingen om sneller te reageren op kwetsbaarheden en incidenten. Voor u betekent dit dat de tijd om te reageren korter wordt. Aanvallers kunnen sneller toeslaan nadat een kwetsbaarheid bekend wordt. Daardoor heeft u minder ruimte om problemen te signaleren en op te lossen.

Waarom is dit belangrijk voor u?

Vooral middelgrote en kleinere ondernemingen lopen extra risico. Bijvoorbeeld omdat de beveiliging minder ver is ontwikkeld of omdat er nog oudere systemen worden gebruikt. Tegelijk kan de impact van cyberincidenten groot zijn, ook in de keten waarin u werkt.

De AFM volgt deze ontwikkelingen nauwgezet en verwacht dat ondernemingen nagaan wat dit betekent voor de eigen organisatie. Wacht daarbij niet af.

Wat kunt u nu doen?

1. Versnel beveiligingsupdates

De tijd om kwetsbaarheden te misbruiken wordt steeds korter. Het is daarom belangrijk om kritieke beveiligingsupdates zo snel mogelijk door te voeren, bij voorkeur geautomatiseerd. Wachten is in de huidige situatie niet meer verstandig.

2. Zorg voor zicht en snelle actie

Zorg dat u weet wat er gebeurt in uw systemen, zodat u afwijkingen tijdig kunt signaleren. U hoeft daarvoor niet direct over geavanceerde oplossingen te beschikken. Ook met basale logging en monitoring, eventueel via uw IT-dienstverlener, kunt u belangrijke stappen zetten.

Zorg daarnaast dat u snel kunt handelen bij incidenten. Beperk schade, bijvoorbeeld door systemen van elkaar te scheiden, het aanvalsoppervlak in kaart te brengen en te zorgen voor betrouwbare back-ups. Test ook regelmatig of uw systemen die via internet bereikbaar zijn voldoende beveiligd zijn.

3. Verken de inzet van AI

AI biedt niet alleen risico's, maar ook kansen. U kunt deze technologie inzetten om kwetsbaarheden sneller te herkennen en te prioriteren. Begin hier gecontroleerd mee en werk waar mogelijk samen met uw leveranciers, met duidelijke afspraken over veiligheid en inzicht. Ook wanneer ICT- of beveiligingsdiensten zijn uitbesteed, blijft het bestuur verantwoordelijk voor de cyberweerbaarheid van die diensten en voor een beheerste aanpak van risico's in de leverancierskeketen.

4. Blijf de basis versterken

Ook in een veranderend dreigingslandschap blijft de basis essentieel, zoals:

- gebruik multi-factor-authenticatie
- beheer toegangsrechten zorgvuldig
- zorg voor inzicht in systemen die via internet bereikbaar zijn

Juist basismaatregelen maken het verschil wanneer aanvallen sneller en eenvoudiger worden uitgevoerd. Voor nadere handvatten verwijzen wij naar de AFM-Principes voor Informatiebeveiliging en de 5 basisprincipes[11] van veilig digitaal ondernemen van het NCSC-NL. Voor ondernemingen die onder DORA vallen, bieden de RTS op het ICT risk management framework[12] daarnaast relevante aanknopingspunten, onder meer voor vulnerability en patch management.

[11] NCSC – 5 basisprincipes digitale weerbaarheid

[12] EUR-Lex – Gedelegeerde Verordening (EU) 2024/1774 / DORA RTS