

Toezichtrapport

Gebruik van AI neemt toe, beheersing en grip blijven achter

CTIVD nr. 86

Vastegesteld op 23 juli 2026

Commissie van Toezicht op de Inlichtingen- en Veiligheidsdiensten

Commissie van Toezicht
op de Inlichtingen- en
Veiligheidsdiensten

CTIVD nr. 86

TOEZICHTRAPPORT

Gebruik van AI neemt toe, beheersing en grip blijven achter

Inhoudsopgave

1. Verkennend onderzoek naar het gebruik van AI en GDA 3
1.1 Aanleiding 3
1.2 Scope 4

2. Bevindingen 6
2.1 Gebruik van AI en GDA door de diensten 6
2.2 AI en GDA-systemen worden centraal en decentraal ontwikkeld en gebruikt 6
2.3 Beleid met passende waarborgen moet nog worden geïmplementeerd 7

3. Conclusies 8

Commissie van Toezicht
op de Inlichtingen- en
Veiligheidsdiensten

CTIVD nr. 86

TOEZICHTRAPPORT

**Gebruik van AI neemt toe,
beheersing en grip blijven achter**

1. Verkennend onderzoek naar het gebruik van AI en GDA

1.1 Aanleiding

Op 30 september 2025 heeft de CTIVD aangekondigd een verkennend onderzoek te gaan doen naar het gebruik van Artificial Intelligence (AI) en andere vormen van Geautomatiseerde Data Analyse (GDA) door de Algemene Inlichtingen- en Veiligheidsdienst (AIVD) en de Militaire Inlichtingen- en Veiligheidsdienst (MIVD) (hierna samen: 'de diensten').

Het doel van dit verkennende onderzoek is om een beter beeld te krijgen in hoeverre AI en GDA door de diensten worden gebruikt en de daarmee gepaard gaande risico's te identificeren. Op basis van de uitkomsten van dit onderzoek kan de CTIVD bepalen hoe het toezicht op dit onderwerp verder moet worden ingericht. De CTIVD oordeelt in dit onderzoek nog niet over de rechtmatigheid van de toepassing van AI en GDA door de diensten. In deze publicatie rapporteert de CTIVD over de uitkomsten van het verkennende onderzoek.

De diensten verwerven en verwerken grote hoeveelheden gegevens om hun wettelijke taken uit te voeren. De diensten verwerken deze gegevens door ze bijvoorbeeld te structureren, filteren of analyseren. Dit is noodzakelijk om van gegevens bruikbare inlichtingen te maken. Tegelijkertijd brengt de verwerking van persoonsgegevens een inbreuk op de privacy van de betrokkenen met zich mee. Dit onderzoek is belangrijk omdat opkomende technologische ontwikkelingen op het gebied van AI en GDA de effectiviteit, maar ook de privacyinbreuk van het verwerken van gegevens kunnen vergroten.

De toepassing van AI blijft toenemen en raakt steeds meer verweven in de samenleving. Er is groeiende aandacht voor de risico's van AI en via verschillende wegen wordt het treffen van waarborgen tegen een aantal negatieve effecten afgedwongen. Zo is bijvoorbeeld binnen de Europese Unie de AI-Act recentelijk in werking getreden en kondigden Nederlandse toezichthouders zoals de Autoriteit Persoonsgegevens aan toezicht te (gaan) houden op de toepassing van AI binnen hun toezichtdomein. Het is de taak van de CTIVD om erop toe te zien dat de diensten gegevens rechtmatig verwerken. Hoewel de AI Act niet van toepassing is op de diensten, biedt de Wet op de Inlichtingen- en Veiligheidsdiensten 2017 (Wiv 2017) een wettelijk kader voor het gebruik van AI en GDA. De Wiv 2017 wordt momenteel herzien. Hiermee beoogt het kabinet onder meer het wettelijk kader voor automatisering aan te passen. De CTIVD wenst met dit rapport een bijdrage te leveren aan deze wetsherziening.

Daarnaast gaven de bevindingen uit het toezichtrapport nr. 74 (2021)[1] en het toezichtrapport nr. 69 (2020) van de CTIVD[2] aanleiding voor het doen van dit verkennende onderzoek. In het eerstgenoemde rapport over Automated Osint (rapport nr. 74) merkt de CTIVD op dat het "*(...) essentieel is dat de diensten voorafgaand aan de concrete inzet weten hoe de tools werken en welke bronnen daarbij kunnen worden geraadpleegd. Alleen met deze kennis kan een gedegen toets worden uitgevoerd om te bepalen hoe de verwerking van deze gegevens met deze tools zich verhoudt tot de algemene bepalingen omtrent gegevensverwerking uit de Wiv 2017*". In het tweede rapport, nr. 69 uit 2020, constateerde de CTIVD dat in het kader van GDA de vertaalslag van wetgeving en beleid naar de uitvoeringspraktijk een aandachtspunt was. Hierin wordt ook benoemd dat veel werkinstructies ontbraken.

1.2 Scope

De CTIVD heeft onderzoek gedaan naar het gebruik van AI en GDA door de diensten. Dit betreft bijvoorbeeld de toepassing ervan in het inlichtingenproces, IT-support of ten behoeve van reguliere bedrijfsvoeringsprocessen.

De definities van AI en GDA die in dit onderzoek zijn gehanteerd, staan beschreven in onderstaande groene kaders.

Artificial Intelligence systeem (AI-systeem)

Een AI-systeem is een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te generen zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen.

(Artikel 3 van de AI Act)

AI kan in verschillende vormen en voor verschillende doeleinden worden gebruikt. Wanneer AI bijvoorbeeld wordt gebruikt om op geautomatiseerde wijze gegevens te analyseren, kan de toepassing van AI worden gezien als GDA.

De diensten zijn bevoegd om GDA toe te passen ten behoeve van de verwerking van gegevens. Artikel 60 lid 2 Wiv 2017 beschrijft een drietal vormen van GDA: gegevens mogen op geautomatiseerde wijze onderling met elkaar (dan wel in combinatie met elkaar) worden vergeleken, doorzocht aan de hand van profielen en vergeleken met het oog op het opsporen van bepaalde patronen.

1 Link naar rapport 74
2 Link naar voortgangsrapportage

Geautomatiseerde data-analyse (GDA)

De wet geeft geen definitie van het begrip GDA en dus ook geen scherpe afbakening van wat wel en niet onder GDA kan worden verstaan. Hierdoor kan ook een simpele zoekslag met een computer als GDA worden gekwalificeerd.

In dit onderzoek heeft de CTIVD zich gericht op toepassingen van GDA met gelijksoortige risico's als die van AI-systemen. Een voorbeeld van een dergelijke GDA-toepassing is een GDA-systeem om gegevens te verwerken waarin de verwerkingsslag dusdanig complex is dat de genomen stappen in het proces en de output niet meer te volgen zijn. Ook bestaan er systemen die zo zijn gemaakt, dat men niet meer naar de onderliggende werking van het systeem kan kijken. Denk hierbij bijvoorbeeld aan complexe rule-engines. Dit zijn niet-lerende algoritmes, waarvan de uitkomsten niet altijd gemakkelijk zijn uit te leggen. Ook is het mogelijk dat de diensten algoritmes inkopen of overnemen van een partnerdienst, waarbij er geen of slechts beperkt zicht is op de werking van het algoritme. Om deze reden ziet het verkennend onderzoek toe op zowel AI als GDA, waarbij de focus ligt op de mogelijke risico's die zich voordoen. De wat eenvoudigere zoekslagen, waarbij dergelijke risico's zich niet of in zeer kleine mate voordoen, vallen niet binnen de scope van dit onderzoek.

Het onderzoek is gericht op AI en GDA waarvan de toepassing risico's voor de rechten en vrijheden van personen met zich meebrengt.

2. Bevindingen

2.1 Gebruik van AI en GDA door de diensten

De diensten maken gebruik van diverse AI- en GDA-systemen. Deze systemen worden voor verschillende doeleinden ontwikkeld en/of gebruikt. Denk aan het samenvatten van teksten, het uitvoeren van netwerkanalyses en het automatisch aanvullen van een gegevensstroom met vertalingen. Deze systemen leveren de dienst veel voordelen op, met name op het gebied van snelheid en efficiëntie van het verwerken van gegevens. Ook biedt het gebruik van AI en GDA meer mogelijkheden voor het leggen van verbanden tussen verschillende gegevens(bronnen) en het analyseren van grote gegevensbestanden zoals bulkdatasets. De hoeveelheid te verwerken gegevens neemt bovendien toe, waardoor verwerking op grotere schaal plaatsvindt.

AI- en GDA-systemen worden binnen de diensten zelf ontwikkeld, volledig of deels verkregen en vervolgens zelf verder ontwikkeld. Het verkrijgen van systemen kan op verschillende manieren, bijvoorbeeld door deze te kopen van een commerciële partij, gratis online aan te schaffen of te verkrijgen van een partnerdienst. De diensten kiezen er in de praktijk voornamelijk voor om systemen intern te ontwikkelen of deels te verkrijgen. Verkregen systemen kunnen worden bijgetraind met interne gegevens of aangepast, zodat deze voldoen aan de specifieke behoeften van de diensten.

Bij het gebruik van AI en GDA kan sprake zijn van risico's op het gebied van onder meer transparantie, betrouwbaarheid en schaalbaarheid. Zo kan een verkregen model getraind zijn op gegevens die niet representatief zijn voor het uiteindelijke gebruik bij de diensten. Hierdoor kunnen de uitkomsten van het model minder betrouwbaar zijn. De CTIVD constateert dat deze risico's zwaar wegen in de beslissing van medewerkers om externe systemen en modellen aan te schaffen. Het is echter onduidelijk of alle medewerkers en teams dezelfde afwegingen betrekken bij dergelijke beslissingen.

Uit het verkennende onderzoek blijkt verder dat niet voor alle medewerkers duidelijk is wanneer ze gebruikmaken van een AI of GDA-systeem. Als gevolg hiervan bestaat het risico dat een gebruiker zich niet bewust is van de beperkingen van zo'n systeem en de kaders voor het gebruik ervan. Hierdoor kunnen systemen verkeerd worden gebruikt of uitkomsten verkeerd worden geïnterpreteerd. De diensten ondernemen actie om bewustwording hieromtrent bij de medewerkers te vergroten. Zo is er aandacht voor AI-geletterdheid en het tonen van waarschuwingen of aanvullende informatie om medewerkers in specifieke gevallen duidelijk te maken dat ze gebruikmaken van een AI- of GDA-systeem. In een aantal gevallen wordt de medewerker geïnformeerd over de betrouwbaarheid van de uitkomst van het systeem.

2.2 AI en GDA-systemen worden centraal en decentraal ontwikkeld en gebruikt

Gedurende het verkennende onderzoek heeft de CTIVD een goed beeld gekregen van de ontwikkeling en het beheer van centrale AI- en GDA-systemen. Binnen de organisatie is een aangewezen afdeling verantwoordelijk voor de ontwikkeling en het beheer van AI- en GDA-systemen. De systemen die worden ontwikkeld of beheerd door deze afdeling worden verder aangeduid als 'centrale systemen'. Tegelijkertijd is gebleken dat de ontwikkeling van AI- en GDA in de praktijk ook elders binnen de diensten plaatsvindt, waarbij afzonderlijke afdelingen, teams of individuele medewerkers zelf systemen kunnen ontwikkelen. Dit wordt in dit rapport aangeduid als 'decentrale systemen'.

Voor de decentrale ontwikkeling en het beheer van AI- en GDA-systemen bestaan verschillende redenen. Zo is sprake van een grotere behoefte aan systemen dan de centrale afdeling kan ontwikkelen (o.a. vanwege toebedeling van capaciteit) en is er behoefte aan team-specifieke functionaliteiten. Daarnaast is de ontwikkeling van nieuwe systemen een tijdsintensief proces, wat ertoe kan leiden dat afdelingen, teams of individuele medewerkers ervoor kiezen om sneller zelfstandig systemen te ontwikkelen met eigen kaders.

Decentrale ontwikkeling van systemen hoeft op zichzelf niet problematisch te zijn, maar de werkwijze zoals deze op dit moment plaatsvindt bij de diensten brengt risico's met zich mee. Uit dit verkennende onderzoek blijkt bijvoorbeeld dat er verschillen zijn tussen werkwijzen en kaders die worden gehanteerd bij het valideren van systemen, afstemming van uitkomsten met collega's, het gebruik van prototypes of het bepalen van de minimale vereisten. De diensten ondernemen wel stappen om risico's te mitigeren. Zo wordt een nieuw algoritmebeleid ontwikkeld en is er aandacht voor het opzetten van een interne registratie van algoritmes. Ondanks de reeds ondernomen stappen ontbreekt het overzicht van en inzicht in de ontwikkeling en het beheer van systemen. Daarnaast is het voor zowel de diensten als voor de CTIVD onduidelijk of alle systemen conform de geldende wettelijke eisen en intern beleid zijn ingericht en worden gebruikt.

2.3 Beleid met passende waarborgen moet nog worden geïmplementeerd

De diensten hebben in 2020 gezamenlijk GDA-beleid vastgesteld. Daarnaast gelden andere overlappende beleidsstukken zoals beleid over het verbod van geautomatiseerde besluitvorming. Kort na de vaststelling van het GDA-beleid werden een aantal deelprojecten gestart om het beleid te implementeren. Uiteindelijk is het beleid niet volledig geïmplementeerd en werd het beleid niet werkbaar bevonden. Het beleid was dusdanig omvangrijk en onvoldoende concreet, dat medewerkers het beleid niet in de praktijk konden toepassen.

Mede naar aanleiding daarvan zijn de diensten in 2024 opnieuw gestart met het opstellen van algoritmebeleid. Dit beleid bevat fundamentele kaders en principes voor de verantwoorde ontwikkeling, inzet en het gebruik van algoritmische systemen binnen de diensten. Het streven van de diensten is om het beleid in 2026 vast te stellen. In het beleid is aandacht voor beginselen als transparantie en AI-geletterdheid. Het beleid bevat verder een robuust pakket aan mitigerende maatregelen en waarborgen, zoals eisen omtrent het gebruik van trainingsdata of controle en validatie van uitkomsten. Ook wordt een werkgroep opgericht die zich zal buigen over operationele risk- en compliance vraagstukken omtrent AI- en GDA-gebruik.

Aangezien het beleid nog niet wordt toegepast, is het treffen van waarborgen en het verminderen van risico's bij de ontwikkeling en het gebruik van AI en GDA sterk afhankelijk van de afdeling(en), teams of individuele medewerkers die de AI- en GDA-systemen ontwikkelen. Daarnaast kunnen aanzienlijke verschillen ontstaan tussen de al dan niet getroffen mitigerende maatregelen en waarborgen.

De CTIVD heeft geconstateerd dat bij de centraal aangewezen afdeling medewerkers werken die zich bewust zijn van de risico's van het gebruik van AI en GDA. De afdeling werkt met eigen procedures en werkafspraken waarbij aandacht is voor beginselen zoals transparantie, betrouwbaarheid en menselijke tussenkomst. Er zijn echter ook voor deze afdeling nog geen werkbare centraal vastgestelde procedures en beleidsstukken voorhanden. Als gevolg hiervan rust op hen een grote verantwoordelijkheid bij het ontwikkelen en beheren van AI- en GDA-systemen.

3. Conclusies

De diensten maken in toenemende mate gebruik van AI en GDA bij de uitvoering van hun wettelijke taken. De systemen worden grotendeels intern ontwikkeld. Daarnaast worden verkregen systemen gebruikt die kunnen worden aangepast en bijgetraind om te voldoen aan de behoeften van de diensten.

De ontwikkeling en het gebruik van de systemen brengen de noodzaak met zich mee om te werken binnen duidelijke kaders, inclusief passende mitigerende maatregelen en waarborgen ter bescherming van rechten en vrijheden van personen. De toepassing van AI en GDA zijn in de praktijk voortgezet, terwijl de daarvoor benodigde kaders en waarborgen niet tijdig en in samenhang zijn ingericht.

De centrale afdeling die verantwoordelijk is voor de ontwikkeling en het beheer van systemen, bestaat uit professionals die zich bewust zijn van de risico's en daarvan blijk geven in hun handelen. Dit neemt niet weg dat ook voor hen heldere, vastgestelde kaders noodzakelijk zijn om rechtmatig en zorgvuldig te kunnen werken. Op hen rust nu een grote verantwoordelijkheid voor het rechtmatig en zorgvuldig ontwikkelen en beheren van AI- en GDA-systemen. Dit creëert een afhankelijkheid van deze individuele medewerkers. Dit is een kwetsbaarheid die verholpen kan worden door centrale inbedding van kaders en werkwijzen.

Voor decentrale ontwikkeling en gebruik van AI en GDA geldt dat de diensten hierop beperkt zicht hebben. De diensten zijn hierdoor onvoldoende 'in control' en ook voor de CTIVD bemoeilijkt dit het verkrijgen van inzicht in mogelijk onrechtmatige toepassing van AI en GDA.

Hoewel de diensten duidelijke stappen zetten en er aantoonbaar aandacht is voor het onderwerp, verloopt de ontwikkeling en toepassing van AI en GDA sneller dan de implementatie van bijbehorende kaders, waarborgen en AI-geletterdheid. Daarnaast hangt het effect van het opgestelde beleid af van de implementatie die nog moet plaatsvinden. Verder is het van belang dat er aandacht blijft voor de werkbaarheid van het beleid en dat de beleidsontwikkeling leidt tot het bedoelde effect. Ondanks de ondernomen stappen zijn de diensten op dit moment nog niet in control.

Het is aannemelijk dat de diensten de komende tijd nog meer AI en GDA willen gaan ontwikkelen en toepassen. Het is daarom van belang te voorkomen dat de reeds opgebouwde achterstand verder oploopt. Hierbij betrekt de CTIVD haar conclusie in de voortgangsrapportage IV van 2020 waaruit bleek dat de vertaalslag van wetgeving en beleid naar de uitvoeringspraktijk een belangrijk aandachtspunt is. Met dit verkennende onderzoek concludeert en benadrukt de CTIVD dat deze vertaalslag in 2026 nog steeds onvoldoende is gerealiseerd.

Samengevat concludeert de CTIVD dat de diensten stappen moeten zetten om te komen tot verantwoord gebruik en ontwikkeling van AI en GDA. De CTIVD gaat met de diensten in gesprek over de wijze waarop zij aan de CTIVD zullen rapporteren over de voortgang en het tijdpad waarbinnen de stappen zullen worden gezet.



Postbus 85556, 2508 CG Den Haag

T 070 315 58 20
E info@ctivd.nl | www.ctivd.nl