# AI en criminaliteit: slimme(re) criminaliteit en slimme(re) criminaliteitsbestrijding?

Brands, J.; Zand, E.G. van 't; Rodermond, E.; Roks, R.

# Criminologie

## *DD* 2024/40

### AI en criminaliteit: slimme(re) criminaliteit en slimme(re) criminaliteitsbestrijding?

#### 1. *Inleiding*

*De opkomst van artificiële intelligentie (AI) markeert een transformatie in diverse industrieën, waarbij technologieën zoals machine learning en neurale netwerken taken automatiseren die voorheen menselijke intelligentie vereisten. Deze vooruitgang heeft geleid tot aanzienlijke verbeteringen in efficiëntie, nauwkeurigheid en innovatie, met toepassingen variërend van medische diagnostiek tot autonome voertuigen. Tegelijkertijd brengt de snelle evolutie van AI ook ethische en maatschappelijke uitdagingen met zich mee, waaronder kwesties rondom privacy, bias, en de impact op de arbeidsmarkt.*

De titel en de daarop volgende drie zinnen zijn niet verzonnen en geschreven door de auteurs van de onderhavige bijdrage, maar zijn het antwoord van Chat GPT op de prompts: '*Schrijf drie zinnen voor een wetenschappelijke tekst over de opkomst van AI'* en '*verzin een leuke, pakkende titel voor een bijdrage over artificiële intelligentie en criminaliteit'*. Het bovenstaande dient ter illustratie van de constatering van de Wetenschappelijke Raad voor het Regeringsbeleid in haar rapport "Opgave AI" dat AI niet zomaar als een technologie moet worden gezien, maar als een "systeemtechnologie die de samenleving fundamenteel zal veranderen" (2021:11). De term AI wordt gebruikt om te verwijzen naar "systemen die intelligent gedrag vertonen door hun omgeving te analyseren en actie ondernemen om specifieke doelen te bereiken" (Schuilenburg, 2024:18). Met het toenemende aantal alledaagse toepassingen van AI, zoals in navigatie of de content die wij als consumenten krijgen voorgeschoteld op streamingsdiensten op basis van ons kijkgedrag, komen ook de nodige zorgen. AI wordt omschreven als onzeker, ongrijpbaar en onvoorspelbaar (Custers et al., 2021) en er wordt gewezen op de risico's wanneer de technologie onze maatschappelijke belangen niet dient, zoals bij de verspreiding van *fake news,* de beïnvloeding van verkiezingen of door de overheid gebruikte AI-systemen met ingebouwde biases zoals in het geval van de toeslagenaffaire.

In een recent themanummer van *Justitiële Verkenningen* over AI, justitie en veiligheid merken Schuilenburg en Soudijn (2024:13) op dat AI nog amper zichtbaar is op de criminologische radar, ondanks de relevantie voor het vakgebied. Allereerst biedt AI nieuwe mogelijkheden tot het plegen van criminaliteit. Al langer wordt erop gewezen dat AI misbruikt kan worden, en ook daadwerkelijk wordt, voor criminele activiteiten. In de wetenschappelijke literatuur worden allerhande tegenwoordige en toekomstige *"crime and security threats associated with AI"* (Caldwell e.a., 2020:2) uiteengezet. Bepaalde auteurs spreken in deze context over het kwaadwillend *("malicious")* gebruik van AI. Wat hieronder verstaan kan worden en welke dreigingen hierbij worden onderscheiden, zullen we in de volgende paragraaf beschrijven op basis van een selectie van de literatuur die sinds 2018 over dit onderwerp is verschenen. Daarnaast zou AI voor de opsporing en bestrijding van criminaliteit een *game changer* (kunnen) zijn, in het bijzonder omdat AI – anders dan de mens – in staat is om enorme hoeveelheden data te analyseren en daarin patronen te herkennen. Op dit onderwerp komen we later in deze bijdrage terug. We besluiten deze bijdrage, die verder overigens exclusief door de auteurs is gegenereerd, met het opwerpen van enkele vragen rondom de ontwikkelingen op het gebied van AI voor de criminologie.

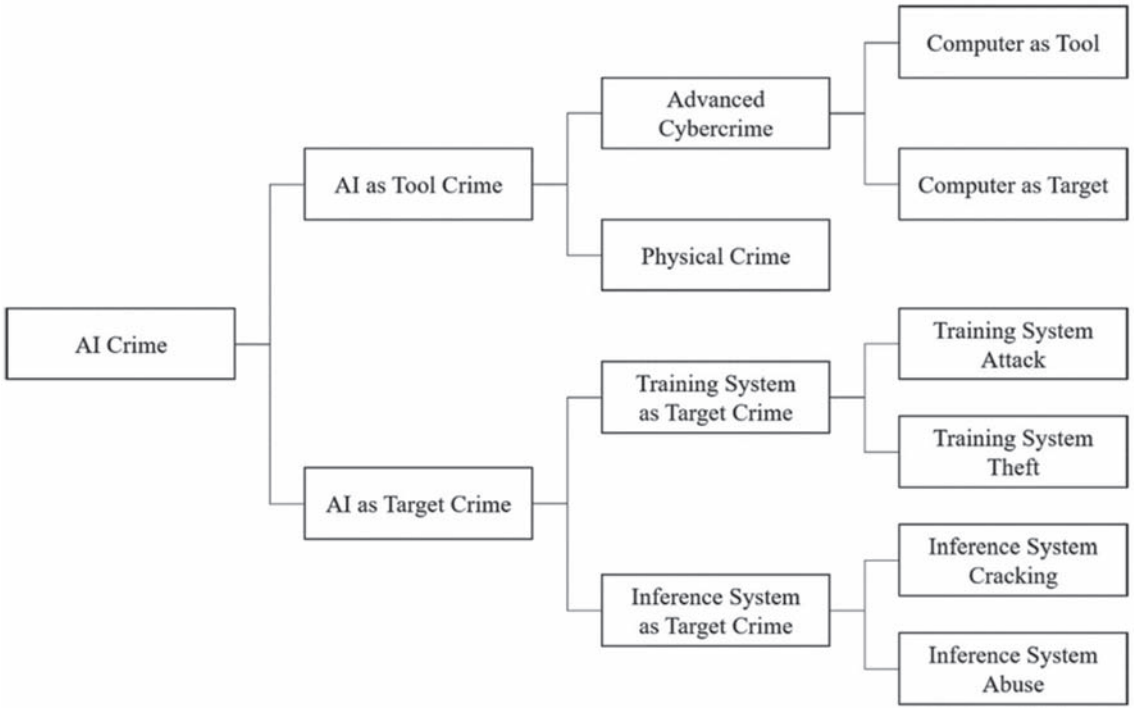
## 2. *Verschijningsvormen en dreigingen van AI-criminaliteit*

In de beschikbare wetenschappelijk literatuur wordt op verschillende manieren gepoogd een overzicht te geven van verschijningsvormen van AI-criminaliteit. Een eerste benadering in de literatuur over criminaliteit en bedreigingen voor de veiligheid die te relateren is aan AI richt zich op bepaalde verschijningsvormen van criminaliteit waarin AI (in de toekomst) een rol kan spelen. Een voorbeeld hiervan vinden we terug in de systematische en interdisciplinaire literatuurreview van King en collega's (2020). Als eerste bespreken zij *"commerce, financial markets, and insolvency"*. Hierbij valt te denken aan zogenoemde "pump-en-dump-praktijken". In deze gevallen laten groeperingen door *bots* massaal valse informatie verspreiden om de prijs van een aandeel 'op te pompen' en deze met winst te verkopen, waarna de waarde van het aandeel in elkaar stort. Bij de tweede categorie die de auteurs onderscheiden, de *"harmful or dangerous drugs"*, wordt gewezen op het gebruik van onbemande vervoermiddelen in het kader van drugssmokkel. Tegelijkertijd wordt ook hier de mogelijke rol van *bots* besproken, waarbij het denkbaar is dat deze gebruikt worden om de verkoop van verdovende middelen te faciliteren.

King en collega's (2020) bespreken vervolgens *"offences against the person"*, waarbij een nadruk ligt op intimidatie. Ook hier kunnen *bots* ingezet worden om mensen te intimideren, zowel direct (schadelijke berichten verspreiden) als indirect (opnieuw delen van andermans schadelijke berichtgeving). De risico's op intimidatie nemen bovendien toe doordat met behulp van AI steeds ingewikkeldere content gecreëerd kan worden, zoals in de vorm van foto's, video's en audio. De koppeling naar de volgende categorie in het overzicht van King en collega's (2020), te weten *"theft and fraud, and forgery and personation"*, is hierna gemakkelijk te leggen. AI kan hier gebruikt worden om enerzijds informatie te verzamelen, bijvoorbeeld door *bots* in te zetten om 'vriendschappen' te sluiten op sociale media netwerken, waarna privacygevoelige informatie beschikbaar wordt. Anderzijds kunnen de aspecten van AI gebruikt worden om (samen met verzamelde persoonlijke informatie) personen te misleiden of af te persen. Ten slotte bespreken King en collega's (2020:104) seksuele misdrijven, waarbij AI *"could be used to facilitate representations of sexual offences, to the extent of blurring reality and fantasy"*. Hoewel deze manier van kijken ons veel leert over de rol die AI speelt – of kan spelen – bij het plegen van criminaliteit, werpt het tegelijkertijd een beperking op. Door op deze manier te kijken blijven eventuele *nieuwe* typen criminaliteit die ontstaan door de intrede van AI buiten beeld.

Andere bijdragen in de literatuur over AI maken dan ook onderscheid naar enerzijds typen criminaliteit waarbij AI "als hulpmiddel wordt gebruikt om traditionele vormen van criminaliteit te plegen" (Caldwell et al., 2020; Hayward en Maas, 2020; Schuilenburg en Soudijn, 2024:16) en anderzijds naar strafbare feiten die gepleegd worden 'tegen AI-systemen' (Hayward en Maas, 2020; Schuilenburg en Soudijn, 2024:19). Jeong maakte in 2020 een vergelijkbaar onderscheid naar *"AI as tool crime"* en *"AI as target crime"* (zie figuur 1).

**Figuur 1: Een taxonomie van AI crime, zoals geformuleerd in en gekopieerd uit Jeong (2020: 184564): DOI: 10.1109/ACCESS.2020. 3029280. Licenced under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.**




Bij de eerste variant wordt AI dus gebruikt als hulpmiddel om bestaande vormen van criminaliteit te plegen. Hierbij kan volgens Jeong (2020) worden onderscheiden naar fysieke criminaliteit en cybercriminaliteit. Bij het eerste kunnen we denken aan toepassingen waarbij voertuigen aangestuurd worden door AI, bijvoorbeeld om fysieke schade aan te richten, of om drugs te smokkelen zoals wij hierboven ook al noemden. Onder cybercriminaliteit deelt Jeong (2020) typen criminaliteit op zoals dat eerder al gedaan is binnen de algemene cybercriminaliteit literatuur; naar *cyber-dependent crime* (waarbij ICT-systemen het doelwit zijn: *"computer as target"*) en gedigitaliseerde criminaliteit of *cyber-enabled crime* (vormen van criminaliteit die los van ICT ook plaatsvinden, maar waarbij ICT als hulpmiddel gebruikt wordt: *"computer as tool"*). Door gebruik te maken van AI kunnen daders het bereik en/of de effectiviteit van beide typen cybercriminaliteit vergroten. Te denken valt aan de mogelijkheden die AI creëert om het te doen laten lijken dat "iemand iets zegt of doet, wat in werkelijkheid nooit gebeurd is" middels beeld en audio (*deepfakes* en *voice cloning*), hetgeen gebruikt kan worden voor "strafbare delicten als kinder- en wraakporno, voyeurisme, misleiding, oplichting en fraude" (Schuilenburg en Soudijn, 2024:16). Ook politieke manipulatie (Hayward en Maas, 2020), en eerder besproken *pump-en-dump schemes* worden in de literatuur besproken als vormen van gedigitaliseerde criminaliteit. Specifiek *cyber-dependent crime* kan volgens Jeong (2020) met behulp van AI gepleegd worden door bijvoorbeeld kwetsbaarheden in ICT-infrastructuren op te sporen.

Bij criminaliteit waarbij AI-systemen het doelwit zijn, maakt Jeong (2020) onderscheid naar het trainings- en het inferentiesysteem. Op basis van data wordt binnen het trainingssysteem een getraind algoritme gecreëerd. Dat algoritme wordt binnen het inferentiesysteem gebruikt om nieuwe data te classificeren (Jeong, 2020: 184564).

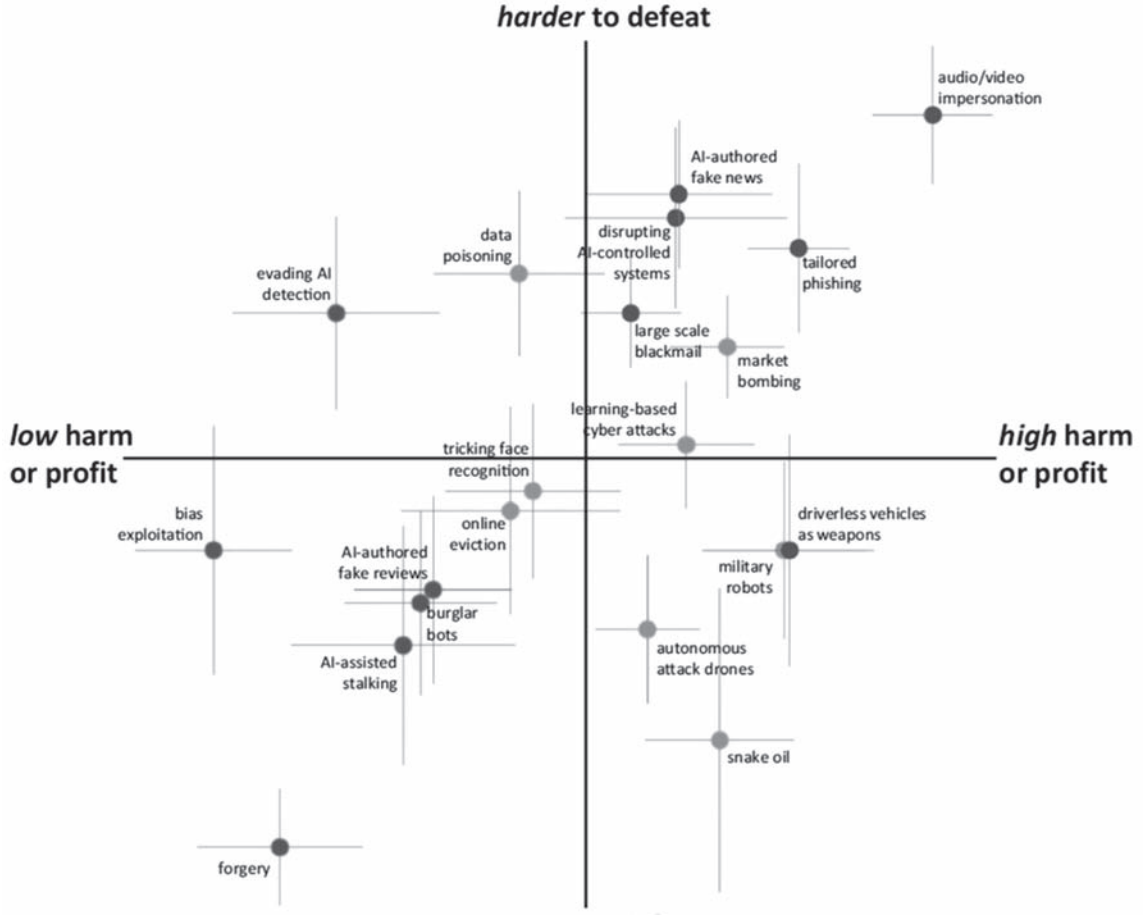
Kwaadwillenden kunnen toevoegingen en/of aanpassingen maken aan het trainingssysteem. Zo bespreken Liu en collega's al in 2018 *"poisoning attacks"* als bedreiging. De publicatie "AI-systemen: ontwikkel ze veilig" van de Algemene Inlichtingen- en Veiligheidsdienst (AIVD) biedt toegankelijke uitleg over deze (en andere, zie hieronder) aanvallen: "Met een poisoning aanval probeert een aanvaller aanpassingen te maken in je data, zodat het AI-systeem wordt 'vergiftigd' en daardoor niet meer werkt zoals gewenst" (AIVD, 2023:7). We kunnen hierbij denken aan foutieve classificaties bij spamfilters of malwaredetectie (Liu e.a., 2018).

Ook het inferentiesysteem kan (op verschillende manieren) aangevallen worden. Bij *evasion* (Liu e.a., 2018), ook wel input-aanvallen, werkt het AI-systeem niet (naar behoren) door bewerkte input. Hierbij geeft de AIVD (2023:17) in haar rapportage het voorbeeld van een verkeersbord dat door het opplakken van een post-it een andere betekenis krijgt en ervoor kan zorgen dat een zelfrijdende auto, waarbij het AI-systeem gebruik maakt van input, ongewenste handelingen uitvoert. Een ander voorbeeld is "model reverse engineering & inversion aanvallen" met als doel "om de dataset te reconstrueren die gebruikt is om jouw model te trainen" (AIVD, 2023:7), bijvoorbeeld omdat deze data interessante gegevens bevat voor de aanvaller, hetzij om intellectueel eigendom te stelen of om zwakheden in het systeem te onderzoeken (AIVD, 2023:7; zie ook Liu e.a., 2018).

Verschillende auteurs onderscheiden nog een derde categorie van AI-criminaliteit: *"crimes by AI (AI as intermediary)"* (Hayward en Maas, 2020; Schuilenburg en Soudijn, 2024). Het zelflerend vermogen van AI kan bijvoorbeeld uitmonden in een artificiële actor die mogelijkheden ontdekt om markten te manipuleren of samen te zweren *("collusion").* Dit kan als ongewenst en onvoorzien gevolg optreden, maar dergelijke artificiële actoren kunnen ook opgevat worden als *"criminal shield/intermediary"* (Hayward en Maas, 2020:217) van een individuele dader dan wel criminele organisatie. Het menselijk handelen is hierbij, met andere woorden, naar de achtergrond verdwenen, hetgeen interessante vragen oproept over de mate waarin AI gezien kan worden als actor (Schuilenburg en Soudijn, 2024:20).

Op basis van bovenstaande bespreking (b) lijkt er nadrukkelijk criminele dreiging uit te gaan van AI. Tegelijkertijd bespreken verschillende auteurs dat we niet enkel moeten kijken naar wat *zou kunnen*, of *wat nu (al) kan* en wat (soms) ook al *gebeurt*. Caldwell en collega's (2020) bieden handvatten om het debat over de dreiging van AI in te richten aan de hand van de risico's – een zogenoemde *"threat rating"* – die gepaard gaan met de verschillende vormen van AI-criminaliteit. Bij hun bespreking van AI-criminaliteit met 31 afgevaardigden (uit de wetenschap, private en publieke sector) onderscheiden zij 18 typen criminaliteit. Caldwell en collega's (2020) hebben alle 18 typen beoordeeld op de (sociale/maatschappelijke) schade *("harm")* en hoeveel profijt (*"criminal profit")* deze kunnen opleveren, hoe gemakkelijk deze typen criminaliteit gepleegd kunnen worden *("achievability")*, maar ook hoe gemakkelijk deze tegengegaan kunnen worden *("defeatability").* Caldwell en collega's (2020) hebben deze 18 type AI-criminaliteit vervolgens geplot in een kwadrant, met de dimensie *harm* en *profit* (van hoog naar laag) op de x-as en *defeatability* (eveneens van hoog naar laag) op de y-as, zie figuur 2.

**Figuur 2: Een kwadrant dat defeatability afzet tegen harmfulness of profitability, zoals geformuleerd in en gekopieerd uit Caldwell en collega's (2020: 11): DOI: https://doi.org/10.1186/s40163-020-00123-8. Licensed under a Creative Commons Attribution 4.0 International License.**




De resultaten van het onderzoek van Caldwell en collega's (2020) maken inzichtelijk dat ingewijden zich vooral zorgen lijken te maken over de dreiging van *1) Audio/video impersonation, 2) Tailored phishing, 3) AI-authored fake news* en *4) Disrupting AI-controlled systems*. Uit dit onderzoek uit 2020 leren we dat volgens deskundigen de meeste dreiging uitgaat van het bewerken van beeld en audio om je voor te doen als iemand anders. AI wordt in de eerste drie dreigingen gebruikt als hulpmiddel. De laatste dreiging, *"disrupting AI-controlled systems"* heeft betrekking op de tweede categorie die wij hierboven bespraken, waarbij strafbare feiten gepleegd worden tegen AI-systemen. Hoe dit precies gebeurt, wordt, anders dan in het artikel van Jeong (2020), in het werk van Caldwell en collega's (2020) echter niet toegelicht. Een vraag, gezien de snelle ontwikkeling van AI, is of dit kwadrant er anno 2024 al anders uitziet, en zo ja, hoe.

## 3. *Detecteren en voorkomen van criminaliteit met behulp van AI*

Naast dat AI misbruikt kan worden voor criminele doeleinden, kan AI ook een bijdrage leveren aan het *detecteren en voorkomen* van criminaliteit. Toepassingen van AI in de publieke dienstverlening breiden in rap tempo uit, waarbij met name het gebruik in het sociale en veiligheidsdomein in het oog springt. Op dit vlak zijn de meeste toepassingen gericht op de opsporing van criminaliteit en op inspectie en handhaving (fraudebestrijding), zoals het voorspellen

van veiligheidsrisico's en het identificeren van gedragspatronen (Hoekstra et al., 2021). Daarnaast gelden procesoptimalisatie en onderhoud, maatwerk, forecasting en beleidsvorming, kennisvergaring en archivering als voorbeelden van de vele andere toepassingsmogelijkheden van AI.

Het gebruik van AI binnen de politie is op dit moment echter nog vooral beperkt tot het efficiënter maken van het bestaande politiewerk, met behulp van AI-toepassingen die zowel retrospectief als voorspellend kunnen zijn, maar ook het politiewerk 'in real time' kunnen ondersteunen (Schuilenburg en Soudijn 2024:12-13). Hoewel het voorspellen van criminaliteit minder centraal lijkt te staan, vormt het (retrospectief of 'in real time') toegankelijk maken en analyseren van *big data* een steeds belangrijker component van de opsporing.

Voorgaande zien we ook terug in een recent – voor Politie en Wetenschap geschreven – literatuuroverzicht over "Kunstmatige intelligentie bij de politie" (Kool et al., 2023), waarin drie vormen van AI-toepassingen worden onderscheiden: a) *"predictive policing"*, b) *"smart policing/ smart surveillance"* en c) *"automated policing"*, welke overigens ook overlap vertonen. Bij *predictive policing* analyseert de politie "met behulp van slimme software historische data (…) met als doel om mogelijke criminele gedragingen in de toekomst te voorspellen en daar tijdig op te anticiperen zodat criminele gedragingen voorkomen kunnen worden" (Kool et al., 2023:20; zie ook Landman, 2024). Het bekendste voorbeeld betreft Criminaliteits Anticipatie Systeem (CAS), waarmee criminaliteit in bepaalde buurten of hotspots kan worden voorspeld. Bij *smart policing/ smart surveillance* gaat het om "het monitoren van verdachte activiteiten in de publieke ruimte of online (cybercrime)" (Kool, et al., 2023:33; zie ook Landman, 2024). Dat monitoren gebeurt bijvoorbeeld met behulp van slimme (bewakings)camera's, al dan niet met gezichtsherkenning. Maar ook sociale media monitoring wordt als voorbeeld beschreven. *Automated policing* betreft nog vooral toekomstmuziek, waarbij gedacht kan worden aan intelligente en autonome "robots en gerobotiseerde systemen die door politieorganisaties worden ingezet" (Kool et al., 2023:48). Voorbeelden die in het stuk genoemd worden zijn de inzet van een robot tijdens de gijzeling bij de Amsterdamse Apple store, en een videovoertuig dat in Dordrecht wordt ingezet.

Aangezien er simpelweg steeds meer omvangrijke volumes aan *big data* voorhanden zijn, ligt het voor de hand met behulp van technologie en nieuwe ontwikkelingen daarbinnen deze data op efficiënte wijze te verzamelen, te verwerken en te analyseren. Naast geografische gegevens over criminaliteit kan hierbij gedacht worden aan het analyseren van bulkgegevens die onderschept worden uit gehackte cryptotelefoons. Op basis van zulke gegevens kunnen dan prognoses gemaakt worden en passende interventies ingericht worden. Internationaal gezien springt het gebruik van AI specifiek voor het online detecteren van bepaalde vormen van criminaliteit in het oog. In de wetenschappelijke literatuur wordt veelvuldig stilgestaan bij het gebruik van AI in het detecteren en voorkomen van online seksueel kindermisbruik en terrorisme (*'counter-terrorism'*), waarbij specifiek de (toekomstige) mogelijkheden van toepassingen als patroonherkenning en het *Internet of Things* worden aangestipt (Ganor, 2021; Singh en Nambiar, 2024). Niettemin is de conclusie van voorgenoemd onderzoek in opdracht van Politie en Wetenschap dat er nog weinig empirisch onderzoek beschikbaar is over het gebruik van AI door de politie, volgens de auteurs niet alleen omdat dit nog in ontwikkeling is, maar ook omdat evaluaties intern blijven en niet met de buitenwereld worden gedeeld, mede om te voorkomen criminelen in de kaart te spelen.

Maar in de literatuur horen we niet enkel een roep om voorgenoemde empirische kennisbasis uit te breiden; ook dient er meer ruimte te zijn voor het onderzoeken van de

juridische en ethische implicaties, bijvoorbeeld via audits of periodieke evaluaties (Kool et al., 2023). Volgens de literatuur bestaan er namelijk nadrukkelijke uitdagingen op het gebied van AI-toepassingen binnen het veiligheidsdomein, bijvoorbeeld ten aanzien van regulering, verantwoording en controleerbaarheid – het zogenoemde 'black box'-probleem (Schuilenburg en Soudijn, 2024). De noodzaak voor een kader voor het wegen van effectiviteit, efficiëntie, rechtmatigheid en rechtvaardigheid is evident. Schuilenburg zegt hierover in zijn oratie "Making surveillance public" (2024) dat vanaf de ontwikkeling van AI toepassingen stilgestaan dient te worden bij publieke waarden en gewenste en ongewenste effecten, omdat technologie niet neutraal is. Doordat ICT-ontwikkelaars steeds grotere invloed hebben bij het ontwerpen van de techniek, is aanvullende aandacht nodig om fundamentele rechten en waarden te waarborgen. Hierbij dient er tevens voor te worden gewaakt dat data niet op onrechtmatige wijze worden gebruikt voor andere doelen dan waarvoor ze oorspronkelijk werden verwerkt (een verbredingsproces dat ook wel wordt aangeduid als *'function creep'*).

In het bijzonder wanneer AI zelf beslissingen neemt buiten de controle van beslissingsbevoegden om, bijvoorbeeld middels zelflerende algoritmen (zie hierboven), komen rechtswaarborgen als evenredigheid en rechtsgelijkheid op het spel te staan. Er kan sprake zijn van cirkelredeneren of 'self-fulfilling prophecies' indien de data waarvan de algoritmen gebruikmaken bevooroordeeld of 'vuil' zijn en, dienovereenkomstig, de uitkomsten dat ook zijn, welke het algoritme vervolgens weer gebruikt om zichzelf te verbeteren (Schuilenburg en Soudijn, 2024:20-21). Voorbeelden op het gebied van '*predictive policing*' zijn de op algoritmen gebaseerde fraude-detectiesystemen van de belastingdienst en DUO, die potentiële fraudeurs detecteerden op basis van factoren als migratieachtergrond en daarmee een discriminerende werking hadden. Ook bij het eerder besproken Criminaliteits Anticipatie Systeem (CAS) geldt het euvel dat de zelflerende algoritmen die werden gebruikt voor patroonherkenning complex en ondoorzichtig waren voor de betrokken professionals (Waardenburg, 2021), wat zou kunnen leiden tot overmatige politiecontroles in bepaalde buurten.

Ten slotte bestaan er belangrijke uitdagingen bij het gebruik van AI en *big data* voor *smart policing/smart surveillance*, aangezien dit vaak verder reikt dan alleen de politieorganisatie (*'below, beyond and above the police';* Schuilenburg, 2024). De politie is immers in steeds grotere mate afhankelijk van technologieën ontworpen of gebruikt door burgers, private (Big Tech) partijen en (internationale) samenwerkingsverbanden. Dit kan gaan om video- en automateriaal vanuit smartphones, beveiligingscamera's, slimme deurbellen, autobeveiliging (Tesla's) en allerlei slimme huishoudelijke apparatuur. Maar ook om het delen van politiedata uit Europol netwerken of van financiele (verdachte) transacties via the Egmont Group of Financial Intelligence Units (FIUs) en om het automatisch uitwisselen van DNA, vingerafdrukken en verkeersgegevens via het Europese Prüm systeem.

De grote potentie van zowel *predictive policing* als *smart policing,* en daarbij het gebruik van *big data* en algoritmen, werpt belangrijke juridische en ethische vragen op over de inzet en impact van dergelijke technologieën (Hirsch Ballin, 2024). Waar juridische kaders tekortschieten komt er (onbedoeld) discretionaire bevoegdheid te liggen bij de ontwikkelaars die gaan over het design van de onderliggende AI-technieken, in het geval van de politie dus bij functies als software- en netwerk ontwikkelaar (Schuilenburg en Soudijn, 2023), terwijl zij zich mogelijk nauwelijks bewust zijn van die verantwoordelijkheid. Bovendien zijn de uiteindelijke beslissingsbevoegde personen weliswaar afhankelijk van de output van de ontwikkelde techniek voor hun formele besluitvorming, maar lang niet altijd voldoende in staat om deze

te doorgronden en duiden (Van Eck, et al., 2018; Klievink, 2021; Van Eck et al., 2022). Zo is het, bijvoorbeeld bij zelflerende algoritmen, niet altijd inzichtelijk op basis van welke gegevens, of biases, bepaalde patronen herkend worden. Vanwege de noodzaak te garanderen dat de fundamentele rechten van individuele burgers worden gewaarborgd – welke volgens TNO-onderzoek momenteel nog 'slechts zijdelings' worden betrokken in de ontwikkeling van AI (Hoekstra et al., 2021) – zijn de juristen dus aan zet, alvorens de toepassingsmogelijkheden van AI in de opsporingspraktijk uit te breiden.

## 4. *Discussie*

Het gebruik van AI wordt gezien als een ontwikkeling die de potentie heeft de samenleving fundamenteel te veranderen. AI zal ook op de criminologie een vergaande invloed hebben. In deze bijdrage hebben we ons gebogen over AI en criminaliteit, zowel de diverse verschijningsvormen en dreigingen van AI-criminaliteit als de manier waarop AI gebruikt wordt om criminaliteit te detecteren en voorkomen. We zien hierbij slimme(re) vormen van criminaliteit die zich voor een deel niet meer beperken tot menselijk handelen. Aangezien hiervoor slimme(re) technologieën worden ingezet, lijkt ook voor de opsporing ervan het gebruik van AI aangewezen te zijn. Het is echter de vraag of de inzet van technologie gezien kan worden – of moet worden – als de oplossing voor uitdagingen die juist mede het gevolg zijn van technologische ontwikkelingen. In de literatuur lijkt deze vraag voor een belangrijk deel nog onbeantwoord. Wel wijzen verschillende auteurs op de mogelijkheid van meer regulering of criminalisering van de als grootste bedreigingen bestempelde malafide toepassingen van AI, zoals de eerder genoemde imitatie van video en audio. De impact hiervan is immers immens: van verkiezingsbeïnvloeding en nepnieuws tot *deepfakes* (Van der Helm, 2021; Van der Sloot en Wagensveld, 2022). Kortom, zowel aan de fenomeenkant als op het gebied van *governance* roept AI-criminaliteit veel meer vragen op dat wij in deze beperkte bijdrage hebben kunnen uiteenzetten of beantwoorden. Voor de criminologie als wetenschappelijke discipline vragen de ontwikkelingen rondom AI-criminaliteit om multi-, inter- en transdisciplinaire vormen van samenwerking in zowel onderzoek als onderwijs.

*J. Brands, E.G. van 't Zand, E. Rodermond & R. Roks*

### *Bronnen*

Algemene Inlichtingen- en Veiligheidsdienst (2023). *AI-systemen: ontwikkel ze veilig*. Den Haag.

Caldwell, M., Andrews, J.T., Tanay, T., & Griffin, L.D. (2020). AI-enabled future crime. *Crime Science*, *9*(1), 1-13.

Custers, B. H. M. (2021). Artificiële intelligentie in het strafrecht: een overzicht van actuele ontwikkelingen. *Computerrecht*, *2021*(4), 330-339.

Custers, B., Dechesne, F., de Graaf, T., Meuwese, A., & Wuisman, I. (2021). Waarom (basis) kennis van AI onontbeerlijk is voor juristen. *Ars Aequi*, *70*(12), 1136-1140.

Eck, van M., Bovens, M.A.P. & S. Zouridis (2018). Algoritmische rechtstoepassing in de democratische rechtsstaat. *Nederlands juristenblad*, 93(40), 3008-3017.

Eck, van M. Hout, Van D. & M. Weijers (2022). 'Olievlek op vlek. De zwarte lijst(en) bij de Belastingdienst', *Nederlands juristenblad*, 1606-1616.

Ganor, B. (2021). Artificial or human: A new era of counterterrorism intelligence? *Studies in Conflict & Terrorism*, 44(7), 605-624.

Hayward, K.J., & Maas, M.M. (2021). Artificial intelligence and crime: A primer for criminologists. *Crime, Media, Culture*, *17*(2), 209-233.

Hirsch Ballin M.F.H. (2024). Geautomatiseerde data-analyse in het nieuwe Wetboek van Strafvordering. *BSb*, (1), p. 4-8.

Hoekstra, M., Chideock, C., Veenstra, A. F. van (2021), *Quickscan AI in publieke dienstverlening II*. TNO.

Huisman, W. (2024). Data-analyse door opsporingsautoriteiten: ontwikkelingen, praktijken en risico's. *BSb*, (1), p. 9-14.
Jeong, D. (2020). Artificial intelligence security threat, crime, and forensics: Taxonomy and open issues. *IEEE Access*, *8*, 184560-184574.
King, T.C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions. *Science and engineering ethics*, *26*, 89-120.
Klievink, B. (2021). *Hollen én stilstaan: hoe data en digitalisering de overheid veranderen*. Universiteit Leiden.
Landman, M. (2024). Samenwerken met politiemachines. Politievakmanschap in het tijdperk van artificiële intelligentie. Justitiele Verkenningen, 2024-1, 30-46.
Liu, Q., Li, P., Zhao, W., Cai, W., Yu, S., & Leung, V. C. (2018). A survey on security threats and defensive techniques of machine learning: A data driven view. *IEEE access*, 6, 12103-12117.
Schuilenburg, M. & Soudijn, M. (2023). Big data policing: The use of big data and algorithms by the Netherlands Police. *Policing: A Journal of Policy and Practice*, 17.
Schuilenburg, M. (2024). *Making Surveillance Public. Why we should be more woke about AI and algorithms.* Den Haag: Eleven Publishing.
Schuilenburg, M. & Soudijn, M. (2024). 'AI-criminaliteit: een verkenning van actuele verschijningsvormen', *Justitiële Verkenningen,* 2024-1, p. 12-29.
Singh, S., & Nambiar, V. (2024). Role of Artificial Intelligence in the Prevention of Online Child Sexual Abuse: A Systematic Review of Literature. *Journal of Applied Security Research*, p. 1-42.
Sloot, van der B., Wagensveld, Y. & B-J. Koops (2021). Deepfakes. *De juridische uitdagingen van een synthetische samenleving*, Tilburg University.
Waardenburg, L. (2021). *Behind the scenes of artificial intelligence. Studying how organizations cope with machine learning in practice*, Alblasserdam: Haveka.
Wetenschappelijke Raad voor het Regeringsbeleid (2023). *Opgave AI. De nieuwe systeemtechnologie*. Den Haag.