



## **Generatieve AI als onderzoekstool binnen de rechtspraak(?)**

mr. J.W. van den Hurk, senior raadsheer hof Den Haag en onderzoeker Kenniscentrum Cybercrime

mr. R.J.A. Klaar, juridisch adviseur Kenniscentrum Cybercrime

**Symposium E-Discovery 2026**

## Introductie

### Kenniscentrum Cybercrime voor de Rechtspraak (KCC)

John Remmerswaal, raadsheer hof Den Haag en coördinator KCC

Jan-Willem van den Hurk, senior raadsheer hof Den Haag en onderzoeker KCC

Roel Klaar, juridisch adviseur/onderzoeker KCC

Contact: [cybercrime@rechtspraak.nl](mailto:cybercrime@rechtspraak.nl)

## Inhoud van deze bijdrage

- GAI en rechtspraak: de stand van zaken
- Probleemstelling
- Gebruiksscenario's (kansen en risico's)
  - *Analyse van een strafdossier*
  - *Verkenning van digitaal-forensische onderzoeksthema's*
- Afronding

## GAI en rechtspraak: de stand van zaken (algemeen)

- Rechters lijken zich in toenemende mate bewust te zijn van het gebruik van GAI door advocaten bij het opstellen van processtukken en de mogelijkheid van hallucinaties (zie bijvoorbeeld [dit nieuwsbericht](#))

## GAI en rechtspraak: de stand van zaken (algemeen)

NOS Nieuws • Zondag, 07:13

### Advocaten krijgen waarschuwing en moeten op cursus na verkeerd gebruik AI



**Stan Hulsen**  
redacteur Tech



Drie advocaten hebben tot nu toe een waarschuwing gekregen omdat zij verkeerd gebruik hebben gemaakt van kunstmatige intelligentie zoals ChatGPT. Twee van die advocaten zijn verplicht op een AI-cursus gestuurd, zegt Wouter Timmermans van de toezichthouder op advocaten tegen de NOS.

Het gaat om zaken waarin advocaten een AI-programma gebruikten om hun argumenten bij de rechter te onderbouwen. In die argumenten stonden verwijzingen naar eerdere uitspraken, maar die bleken over iets heel anders te gaan of zelfs helemaal niet te bestaan.

## GAI en rechtspraak: de stand van zaken (juridisch kader)

- De AI-Act houdt geen verbod op het gebruik van GAI-toepassingen door de rechtspraak in, maar de aard van de inzet is bepalend voor de risico-classificatie en (voor zover de inzet niet verboden is) compliance-eisen
- Voor zover GAI niet voor feitenvaststelling of voor toepassing van juridische criteria op feiten wordt ingezet en de inzet geen significant risico op schade oplevert, is er geen sprake van 'hoog risico' en gelden lichtere compliance-eisen
- Tot die compliance-eisen behoren o.a. de GPAI-verplichtingen en transparantie-verplichtingen

## GAI en rechtspraak: de stand van zaken (RechtspraakGPT)

- De **Raad voor de rechtspraak** verbiedt het gebruik van commerciële AI-tools (ChatGPT, Gemini, Claude, etc.) door rechters en secretarissen
- **RechtspraakGPT** ([release](#): 10/12/2025) is een voor de rechtspraak ontwikkelde AI-assistent voor het uitvoeren van tekstuele taken buiten het primaire proces (teksten structureren, samenvatten, herschrijven, etc.)
- RechtspraakGPT maakt gebruik van GPT-4o mini (knowledge cut-off: 1 juni 2024) en is afgesloten van het word-wide-web

## Probleemstelling

***Welke kansen biedt generatieve AI (LLM's) voor de optimalisatie van het dagelijkse werk van rechters, en welke risico's brengt het gebruik mee voor de kwaliteit van de rechterlijke oordeelsvorming en (daarmee) de rechterlijke autonomie?***

## Gebruiksscenario's – analyse van een strafdossier (kansen)

### Structuur van een strafdossier in een complexe(re) (cyber-)zaak

1. Formele stukken (o.a. zaaksinformatie, tenlastelegging, VH-stukken, etc.);
2. Startinformatie / BOB-dossier (verslaglegging opsporingshandelingen);
3. Beslagstukken (o.a. KVI's, processen-verbaal van onderzoek aan gegevensdragers, etc.);
4. Klassiek bewijs (o.a. histo's, transacties, verklaringen van de verdachte);
5. Forensische (NFI-)rapportages
6. **Datasets (o.a. tijdlijnen, CSV-bestanden, logbestanden, etc.)**
7. Stukken OM/verdediging (voor zover toegevoegd aan dossier)

Omvang: in de regel vanaf 1.000 blz.

## Gebruiksscenario's – analyse van een strafdossier (kansen)

- Rechters kunnen een LLM gebruiken om het dossier gericht op zoektermen en/of verbanden bevragen, maar ook voor geavanceerde vormen van zoeken:
  - tijdlijn van gebeurtenissen genereren;
  - hypothesen over de *feitelijke* rollen van de betrokken personen formuleren;
  - een coherent overzicht van alle chatgesprekken in het dossier opstellen en logische verbanden in tijd en plaats herkennen;
  - digitale onderzoeksbevindingen relateren aan niet-digitale bevindingen etc..

## Gebruiksscenario's – analyse van een strafdossier (risico's)

- Een LLM berekent statistische samenhang tussen woorden en concepten
- Een LLM heeft anders dan een rechter geen werkelijk semantisch begrip en daarvan moeten rechters (en andere gebruikers) zich bewust zijn en blijven
- Deze werking van LLM's brengt voor de kwaliteit van rechterlijke oordeelsvorming en voor de rechterlijke autonomie specifieke risico's mee
- De rechter kan niet zelf controleren in hoeverre die risico's in een concreet geval aan de orde zijn

## Gebruiksscenario's – analyse van een strafdossier (risico's)

Werking LLM	Risico's voor kwaliteit van rechterlijke oordeelsvorming
De interactie tussen <b>prompt</b> en het <b>LLM</b> bepaalt welke informatie relevant wordt geacht en wordt verwerkt	Hoe specifieker de vraag, hoe smaller de selectie van de feiten  Geen correctie van (impliciete) aannames in het prompt door het LLM
<b>Onzichtbare vooroordelen</b> in het <b>LLM</b> zelf (algoritmische bias)	Voor de rechter zijn statistische vooroordelen die het LLM op basis van de trainingsdata heeft ontwikkeld niet kenbaar, tenzij het in de output wordt herkend

## Gebruiksscenario's – analyse van een strafdossier (risico's)

Werking LLM	Risico's voor kwaliteit van rechterlijke oordeelsvorming
<p>Fragmentatie van context (<b>chunking</b>) door een beperkt contextvenster (een dossier van &gt; 1000 blz. is lastig om in 1 interactie te verwerken)</p>	<p>Het LLM toont alleen een selectie van relevante fragmenten uit het dossier</p> <p>Het LLM construeert vervolgens een samenhang tussen die voorgeselecteerde relevante fragmenten</p> <p>De rechter ziet niet welke delen van het dossier het LLM buiten beschouwing heeft gelaten</p>

## Gebruiksscenario's – analyse van een strafdossier (hoe moet het **niet**?)

- Een aantal componenten van de beoordeling van strafzaken zou **in geen geval** mogen worden uitbesteed aan een LLM:
  - Op basis van de tenlastelegging en/of verweren bewijsthema's (opzet, schuld, medeplegen, etc.) identificeren
  - Feiten vaststellen
  - Een bewijsmotivering genereren
  - Een strafoplegging in een concrete zaak genereren
- Uitbesteding van deze componenten aan een LLM doet afbreuk aan de rechterlijke autonomie en de kwaliteit van rechterlijke oordeelsvorming

## Gebruiksscenario's – analyse van een strafdossier (hoe moet het **niet**?)

- Een LLM een voorselectie van voor de beoordeling relevante feiten laten maken en het dossier alleen op basis van die voorselectie analyseren
- Een LLM op basis van de tenlastelegging een selectie van relevante juridische criteria (wettelijke voorschriften, jurisprudentie, etc.) laten maken en alleen dat juridische kader tot uitgangspunt nemen voor de beoordeling
- Geen evaluatieve termen ('onredelijk', 'aanmerkelijk', etc.) of normatieve termen ('opzet', 'schuld', etc.) in prompts, omdat het LLM daardoor feiten selecteert die passen bij die termen

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

- Een LLM zou slechts als een hulpmiddel voor het suggereren van zoekrichtingen/hints op basis van (digitale) onderzoeksbevindingen in het dossier moeten worden gebruikt
- Wij onderscheiden hoofdzakelijk 3 mogelijke toepassingen:
  - Het structureren van *alle* voor de beoordeling (mogelijk) relevante (verborgen) verbanden tussen gegevens (**geavanceerd zoeken**)
  - Het vergelijken van specifieke data in 2 (of meer) datasets (**analyse van ruwe data**)
  - Het blootleggen van impliciete aannames in digitale onderzoeksbevindingen (**expliciteren van aannames**)

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

### Voorbeeld 1 (geavanceerd zoeken)

Fenomeen	Voorbeeld van een zoekprompt
Netwerkstructuur	Breng alle in het dossier voorkomende accounts, devices, IP-adressen, telefoonnummers, device ID's en andere metadata) die bij meer dan één persoon voorkomen of met meer dan één persoon in verband worden gebracht, in kaart.
Gebruik van devices	Breng alle in het dossier genoemde IMEI-nummers, serienummers, MAC-adressen en device ID's in kaart.
Aanwezigheid van de verdachte op of binnen een straal van 500 meter van locatie X	Breng alle in het dossier beschreven momenten in kaart waarop een device of account dat aan verdachte wordt toegeschreven, is geregistreerd op of binnen een straal van 500 meter van locatie X.

## Gebbruiksscenario's – analyse van een strafdossier (hoe dan **wel?**)

### Voorbeeld 2 (analyse van ruwe data)

LLM-invoer:

- CSV-bestand met (beweerdelijk) frauduleuze transacties
- Lijst van betalingskenmerken gegenereerd door malafide configuratiebestanden van banking malware

Prompt: “identificeer transacties die malafide betalingskenmerken bevatten, zonder externe aannames en/of feiten in de analyse te betrekken”

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

### Voorbeeld 3 (expliciteren van aannames)

LLM-invoer:

- Digitale onderzoeksbevinding (“het overeen komen van de ‘last modified date’ van de thumbnail met de ‘last modified date’ van de originele afbeelding, betekent dat de gebruiker de originele afbeelding heeft verwijderd”)

Prompt: “identificeer de impliciete aannames waarop deze bevinding berust, zonder dat er feiten aan de bevinding worden toegevoegd”

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

- Betoogd is wel dat voor zover de opvolgende onderzoeksstappen herleidbaar, transparant en reproduceerbaar zijn, en de (digitale) onderzoeksbevindingen en onderliggende digitale sporen controleerbaar zijn, de LLM-hint buiten beschouwing kan blijven (**Henseler & Aben**, [EeR 2025-5](#), redactioneel)
- Deze opvatting lijkt valide voor wat betreft de verantwoording van de bewijsbeslissing (bewijsmotivering), maar niet (zonder meer) voor de rechterlijke oordeelsvormingsprocessen die uitmonden in de bewijsbeslissing
- Het is onzes inziens noodzakelijk om ook procesmatige randvoorwaarden te stellen die oneigenlijke beïnvloeding/sturing van rechterlijke oordeelsvorming mitigeren

## Gebruiksscenario's – analyse van een strafdossier (hoe dan **wel?**)

- LLM-hints vormen een systematisch risico voor de autonomie van de rechter omdat zij de rechterlijke oordeelsvorming oneigenlijk kunnen beïnvloeden:
  - Hints dienen als referentiepunt voor de oordeelsvorming (ankering, en bij routinematig gebruik: cognitieve lock-in)
  - Buiten de hints om komen andere feiten minder snel in beeld (selectiviteit)
  - Hints activeren (onbedoeld) bewijsthema's/rechtsvragen en de oordeelsvorming beweegt zich alleen binnen dat kader (framing)
  - Overschatting van de relevantie en/of betrouwbaarheid van LLM-hints (automation bias, illusie van coherentie/samenhang)

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

- Vanuit het perspectief van rechterlijke autonomie en rechterlijke oordeelsvorming is het gebruik van LLM-hints onzes inziens verantwoord als (in elk geval) aan de volgende randvoorwaarden is voldaan:
  1. Rechters lezen eerst zelf het dossier, maken een voorlopige analyse en gebruiken (eventueel) pas daarna LLM-hints (vermindering van framing en ankering);
  2. Een LLM-hint wordt verplicht onderworpen aan een LLM-tegenlezing (vermindering van bevestigingsdrang);
  3. Bij het volgen van een LLM-hint moet in raadkamer duidelijk zijn waarom de LLM-tegenlezing wordt verworpen (vermindering van bevestigingsdrang, afdwingen van deliberatie).

## Gebruiksscenario's – analyse van een strafdossier (hoe dan **wel?**)

### LLM-tegenlezing (randvoorwaarden)

- Het doel is het problematiseren van de zoekrichting(en) van de LLM-hint
- Andere persona *kan* leiden tot een verandering van gekozen perspectieven en stijl
- Andere LLM gebruiken *kan* tot een betrouwbaardere tegenlezing leiden
- Expliciet verzoeken om ontbrekende informatie (gemiste scenario's, zeldzame maar relevante scenario's, impliciete aannames, etc.)
- Expliciete instructie om resterende onzekerheid te benoemen

## Gebruiksscenario's – analyse van een strafdossier (hoe dan wel?)

### LLM-tegenlezing – voorbeeld 'tegenlees'-prompt

- LLM-hint: De 'last modified date' van kp-afbeelding [X.jpg] betref 15 januari 2024; 00:00. Ik, verbalisant, acht aannemelijk dat de afbeelding toen voor het laatst door de gebruiker is bewerkt." (p. 913 van het dossier)
- Neem de rol aan van een onafhankelijk digitaal-forensisch expert aan met een uitstekend vermogen tot het kritisch analyseren van digitale onderzoeksbevindingen van de politie
- Identificeer ontbrekende informatie die relevant is voor de getrokken conclusie
- Genereer scenario's waarin de registratie van het tijdstempel geen verband houdt met gebruikershandelingen, met bijzondere aandacht voor minder dominante scenario's
- Benoem de resterende onzekerheid

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (kansen)

- In strafzaken zijn regelmatig digitale onderzoeksthema's aan de orde, met name:
  - **Gebruikersidentificatie** (koppeling van devices en/of data aan de verdachte)
  - **Activiteitonderzoek** (koppeling tussen digitale sporen en gebruikershandelingen)
  - **Herkomst van data** (wat is de bron van data)
  - **Integriteit en authenticiteit van data** (hoe betrouwbaar zijn de data)

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (kansen)

### Voorbeelden van onderzoeksvragen

#### Hoe waarschijnlijk is het dat:

- de afbeeldingen in de DCIM-map op de smartphone van verdachte met de camera van diezelfde smartphone zijn vastgelegd?;
- de strafbare gegevens automatisch (zonder voorafgaande gebruikershandeling) op de smartphone van de verdachte terecht zijn gekomen?;
- de registratie van het Windows-tijdstempel 'file created date' verband houdt met opslag van een bestand door een gebruiker van het device?

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (kansen)

- LLM's kunnen voor de rechter een hulpmiddel zijn om:
  - Een basaal begrip van het onderzoeksthema verkrijgen;
  - Mogelijk relevante digitale aspecten van een onderzoeksthema in beeld te krijgen;
  - Op zitting gerichtere vragen over mogelijk relevante digitale aspecten te kunnen stellen;
  - Beter te kunnen anticiperen op (eventuele) technisch-complexe verwerpen op zitting;
  - Voor te bereiden op een forensisch intake (FIT) gesprek, waarin door de r(h)-c nadere afstemming over de formulering van de onderzoeksoopdracht met advocatuur/OM/NFI-deskundigen plaatsvindt;
  - Effectiever met NFI-deskundigen te kunnen communiceren.

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (risico's)

- **Focus op dominante technische scenario's**
  - In LLM-output zullen eerder onderzochte en (in de trainingsdata) dominante technische scenario's worden benadrukt ('majority bias')
  - Nieuwe en/of minder frequent voorkomende technische scenario's blijven onbenoemd of (relatief) onderbelicht, terwijl dat niet kenbaar is voor de rechter
  - Dat kan problematisch zijn als een door het LLM onbenoemd gelaten technisch scenario in het concrete geval wel degelijk relevant blijkt te zijn *en* dat technisch scenario noch door de politie, noch door de NFI-deskundige is onderkend

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (risico's)

- **Focus op dominante technische scenario's – voorbeeld (afwijkende systeemtijd)**
  - Prompt: “formuleer de meest relevante scenario's waarin een afwijking tussen systeemtijd en werkelijke tijd op een Windows-computer door bewuste manipulatie door de gebruiker is veroorzaakt”
  - LLM-output: handmatig klok terugzetten en/of tijdzone wijzigen, handmatige wijziging via command line, automatische wijziging via geplande taak
  - Niet benoemd gebruiker-gerelateerd scenario: configuratie van een valse NTP-server

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (risico's)

- **Illusie van (werkelijk) technisch begrip**
  - De validiteit van digitale onderzoeksbevindingen is steeds afhankelijk van onderliggende aannames en randvoorwaarden waarvan (ook) rechters (met bovengemiddelde digitale kennis) geen compleet beeld hebben
  - Die aannames en randvoorwaarden zullen dus ook niet door de rechter in het prompt kunnen worden geëxpliciteerd
  - Als de rechter zich baseert op een algemene uitleg van een LLM die onvoldoende specifiek is voor het concrete geval, kan dat leiden tot formulering van irrelevante hypothesen

## Gebruiksscenario's – verkenning van digitaal-forensische onderzoeksthema's (risico's)

- **Illusie van (werkelijk) technisch begrip – voorbeeld**
  - Prompt: “wat betekent de ‘file created date’ (FCD) van een Windows-bestand?”
  - LLM-output: “de FCD geeft het exacte moment aan waarop een Windows-bestand voor het eerst op een device is gegenereerd en betreft een sterke aanwijzing voor de opslag van dat bestand door een gebruiker”
  - Rechter: “we kunnen de FCD dus gebruiken om de start van de pleegperiode te markeren en het verweer dat sprake was van automatische opslag onbesproken laten.”

## Afronding ('takeaways')

- De rechter zou eerst zelf (globaal) het dossier moet lezen en *kan* daarna een LLM gebruiken voor het genereren van zoekrichtingen/hints en voor hypothesevorming
- LLM-hints zouden moeten worden onderworpen aan een LLM-tegenlezing en de resultaten daarvan zouden in raadkamer moeten worden besproken
- De rechter dient zich bewust te zijn en blijven van het risico van illusie van begrip dat door het gebruik van LLM's voor hypothesevorming kan ontstaan