Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid



Research & Development en Innovatie Portfolio NFI 2025

Large Language Models (LLM's) en Data Science



Voorwoord

Geachte lezer,

Met trots presenteer ik de laatste editie van het Research & Development (R&D) en Innovatie Portfolio van het Nederlands Forensisch Instituut (NFI). Een bloemlezing van projecten waarmee het NFI bijdraagt aan een toekomstbestendige waarheidsvinding in de strafrechtketen. Stuk voor stuk inspanningen die essentieel zijn voor de forensische doorbraken in toekomstige strafzaken.

Iedereen ondervindt, in zowel werk als privé, de effecten van een steeds meer gedigitaliseerde en datagedreven wereld. De strafrechtketen is daar zeker geen uitzondering op. Kunstmatige intelligentie en alle verschijningsvormen daarvan brengen grote veranderingen met zich mee. Denk bijvoorbeeld aan Large Language Models (LLM's) zoals ChatGPT, die sinds een paar jaar steeds vaker het licht zien. Welke kansen en bedreigingen brengen die met zich mee voor de wereld van forensisch onderzoek?

Als forensisch onderzoeks- en Rijkskennisinstituut gaat het NFI deze uitdagingen aan met behulp van data science. Data science biedt de kans op een verantwoorde manier de mogelijkheden van de technologie van LLM's te vertalen naar alle kennisgebieden die we in huis hebben. Een grote, maar belangrijke uitdaging. Het is dan ook niet voor niets dat 'AI en data science in de forensische praktijk' een thema is in onze Strategische Kennis- en Innovatieagenda (SKIA). De resultaten van onze innovatie inspanningen, vaak met samenwerkingspartners, vindt u in dit R&D- en innovatieportfolio.

Veel leesplezier gewenst,

Annemieke de Vries
Directeur Wetenschap en Technologie
Nederlands Forensisch Instituut



Inhoudsopgave

Voorwoord **2**

Inhoudsopgave **3**

Inleiding **4**

Thema 1 Sporen vinden, veiligstellen en volgen **5**

- Scenariogericht PD-onderzoek (SGPDO)
- Speech-to-text
- Robotiseer referentie-experiment
- Hansken Copilot

Thema 2 De ultieme forensische reconstructie **7**

- Fractuurstudie
- Verklaarbare detectie van synthetisch spraakmateriaal
- 3D-reconstructie chips met SEM
- Objectivering morfologisch onderzoek

Thema 3 Iedereen, overal, alles, altijd **9**

- ForMAT
- Automatische tabletherkenning
- Explosieven in woonwijken - Verborgen verbanden onthullen

Thema 4 AI & data science in de forensische praktijk **11**

- Detector
- MPS vanuit niet-humane matrices
- Agnostische deepfake-detectie
- Hansken RAG
- Wachtwoordanalyse op schaal
- AI voor DNA-analyse

Thema 5 Ketenbreed meer met forensische informatie **14**

- Forensische analyse van geldstromen in ondermijnende criminaliteit
- DNA Beslisondersteuning
- From anonymous to identified

Toekomstperspectief **16**

Bijlage I Projectdetails **17**

Inleiding

Bij het Nederlands Forensisch Instituut beantwoorden we forensische vragen naar aanleiding van een grote diversiteit in strafzaken. Daarmee dragen we elke dag bij aan de waarheidsvinding. Veel van die forensische vragen konden we vijftien jaar geleden minder goed, of zelfs helemaal niet beantwoorden. Omdat we structureel investeren in nieuwe kennis en technieken lukt dat vandaag de dag wel.

Om tijdig te weten welke forensische vragen in de toekomst relevant zullen zijn, kijken we bij het NFI naar de algemene behoefte vanuit de strafrechtketen en vertalen we maatschappelijke, wetenschappelijke en technologische ontwikkelingen naar kansen en bedreigingen voor het strafrecht. Dit helpt ons richting te geven aan onze Research & Development (R&D) activiteiten en aan de innovaties die we uitvoeren samen met kennispartners en de strafrechtketen.

In dit portfolio zijn voorbeelden opgenomen van projecten die naar verwachting gaan bijdragen aan het beantwoorden van de forensische vragen van morgen. Digitalisering en de toepassing van kunstmatige intelligentie neemt al enige tijd een grote vlucht. Dit biedt uitdagingen evenals kansen voor het werk van de forensische keten, waarmee data science onveranderd een grote plek inneemt in het innovatieprogramma van het NFI. De nadruk in dit portfolio ligt dan ook op projecten waarin data science, zoals de toepassing van kunstmatige intelligentie en LLM's, een belangrijke rol speelt. De ongeveer twintig voorbeelden van innovatieprojecten zijn in dit portfolio geordend langs de vijf thema's van de meerjarige Nederlandse Forensische Onderzoeksagenda (NFOA).

Er zijn nog maar weinig forensische onderzoeken waarin geen digitale sporen worden veiliggesteld. Vaak is zelfs sprake van heel véél digitale sporen. Ook het NFI genereert zelf veel data uit de analyses die het uitvoert aan onderzoeksmateriaal. Dit betekent: steeds meer informatie uit steeds kleinere sporen. Data science biedt de mogelijkheid om uit al deze data bruikbare informatie te halen en de informatievoorziening te versnellen. Bij het NFI zetten we data science ook in om nieuwe soorten informatie te kunnen leveren en deze in nieuwe vormen aan te kunnen bieden. Daarmee draagt het NFI bij aan de kansen en uitdagingen van kunstmatige intelligentie in de strafrechtketen, vanuit de rol van forensisch onderzoeks- en Rijkskennisinstituut.

Thema 1
Sporen vinden, veiligstellen en volgen



"Essentieel voor het onderzoeken van misdrijven is het nauwkeurig vinden, veiligstellen en volgen van zelfs het kleinste forensische spoor. Ondersteund door geavanceerde technologie en wetenschap kan het onderzoek op de plaats delict maximaal relevante informatie halen uit minimaal sporenmateriaal. Dit begint met het ontdekken van cruciaal fysiek en digitaal bewijs, waarna het veiligstellen en onderzoeken ervan met zorg moet worden uitgevoerd en vastgelegd."
– NFOA 2023-2033



Sporen vinden, veiligstellen en volgen

Scenariogericht PD-onderzoek (SGPDO)

Het forensisch onderzoek op de plaats delict (PD) is van cruciaal belang voor veel strafrechtelijke onderzoeken. Sporen en informatie die tijdens het opsporingsproces zijn verzameld, worden gebruikt bij de reconstructie en de bewijsvoering van het vermeende misdrijf.
Echter, sporen die niet zijn opgemerkt, niet zijn geselecteerd, niet zijn veiliggesteld of niet goed zijn gedocumenteerd, kunnen weinig tot niets bijdragen aan een strafrechtelijk onderzoek. Momenteel hebben PD-onderzoekers weinig handvatten die hen helpen bij het nemen van de benodigde denkstappen en het maken van de juiste keuzes. In samenwerking met ketenpartners is daarom een praktisch denkkader ontwikkeld, dat PD-onderzoekers ondersteunt in het gericht overwegen van verschillende scenario's en het onderzoeken en veiligstellen van relevante sporen.

Speech-to-text

Het forensisch sporenonderzoek op het laboratorium is een complex proces waarbij goede registratie van waarnemingen en handelingen van groot belang is. Dit project richt zich op het versnellen en verbeteren van deze werkprocessen binnen diverse deskundigheidsgebieden door passende speech-to-text toepassingen te ontwikkelen. Met *near real-time* spraakopname, transcriptie en verwerking met een lokale LLM kunnen registraties sneller en makkelijker verlopen. De ontwikkeling van een *proof-of-practice* heeft veelbelovende resultaten opgeleverd. In de volgende fase zal met een pilot worden onderzocht of deze technologie in de praktijk van meerwaarde kan zijn.

Robotiseer referentie-experiment

Voor data-analyse zijn sporen uit het gebruik van apps nodig als referentiemateriaal in zaken. Er zijn heel veel apps en het handmatige werk om die apps te gebruiken en de sporen vast te leggen is tijdrovend. Het gebruiken van een app willen we automatiseren door het ontwikkelen van een AI referentie-experiment agent op basis van LLM's die zijn getraind om telefoons en computers te bedienen. De meerwaarde hier is: Sneller en efficiënter meer referentiemateriaal maken, wat zorgt voor beter en sneller digitaal forensisch onderzoek. Bovendien kan deze innovatie ter inspiratie dienen voor meer automatisering met AI binnen de strafrechtketen.

Hansken Copilot

Binnen digitaal forensisch onderzoek groeit de hoeveelheid data sneller dan onderzoekers handmatig kunnen analyseren. Het project *Hansken Copilot* richt zich daarom op de ontwikkeling van een AI-ondersteunde assistent binnen Hansken, die onderzoekers helpt bij het samenvatten, doorzoeken en interpreteren van digitale sporen. In 2025 lag de nadruk op experimenten met LLM's die kunnen redeneren over grote hoeveelheden context en daarbij ook gebruikmaken van tools en traceerbare bronverwijzingen. *Hansken Copilot* wordt ontwikkeld als ondersteunend instrument, niet als vervanging van de forensisch onderzoeker. Transparantie, herleidbaarheid en controleerbaarheid staan centraal, zodat de uitkomsten bruikbaar blijven in een forensische en juridische context. Daarmee draagt het project bij aan efficiënter, consistenter en beter uitlegbaar digitaal onderzoek.

Thema 2
De ultieme forensische reconstructie



"Forensisch onderzoek draait om de nauwkeurige reconstructie van strafbare menselijke handelingen. Dit vereist een interdisciplinaire onderzoeksaanpak, interpretatie van het bewijs op activiteitniveau en een gedegen statistische analyse. Experts moeten de forensische reconstructie begrijpelijk kunnen toelichten zonder afbreuk te doen aan de wetenschappelijke kernwaarden."
– NFOA 2023-2033



De ultieme forensische reconstructie

Fractuurstudie

Het project *Fracture Recognition and Analysis in Children: Trauma, Uncommon Genetic Disorders, and Rare Etiologies* onderzoekt of letselpatronen bruikbaar zijn binnen de differentiële diagnostiek tussen niet-accidenteel letsel en ziekten bij minderjarigen. Vooral binnen de Nederlandse pediatrische populatie ontbreekt systematische en vergelijkende data. Met behulp van LLM's zullen systematische reviews een beeld scheppen van de letselpatronen die al beschreven zijn (LLM's helpen bij literatuur screenen, data extractie, en *risk of bias assessments*), zal een database voor letselpatronen beschreven in NFI-rapporten een beeld geven van de geobserveerde letsels de Nederlandse populatie. Dit project biedt de eerste systematische vergelijking, op basis van een cohort, meet de diagnostische waarde van specifieke combinaties van verwondingen en toetst de betrouwbaarheid en meerwaarde van taalmodellen (LLM's) in systematisch onderzoek.

Verklaarbare detectie van synthetisch spraakmateriaal

Synthetische of gemanipuleerde spraak (ook deepfakes; spoofed speech; presentation attacks genoemd) stellen ongekende uitdagingen voor forensisch onderzoek en rechtssystemen. De manier waarop dit type media wordt onderzocht, omvat meestal een black-box classificatie, die ook een waarschijnlijkheid bevat of een spraakfragment echt of nep is. Het doel van dit project is een workflow vaststellen voor de analyse van mogelijk synthetische of gemanipuleerde spraak, die onafhankelijk is van het gebruikte generatieproces. Dit wordt bereikt door het analyseren van spraakkenmerken gebaseerd op o.a. akoestisch-fonetische en taalkundige kennis, die uitlegbaar zijn zowel voor deskundigen als voor leken. Dit wordt ook bereikt door gebruik te maken van zowel shallow- als deep learning-methoden in de detectie van synthetische spraak.

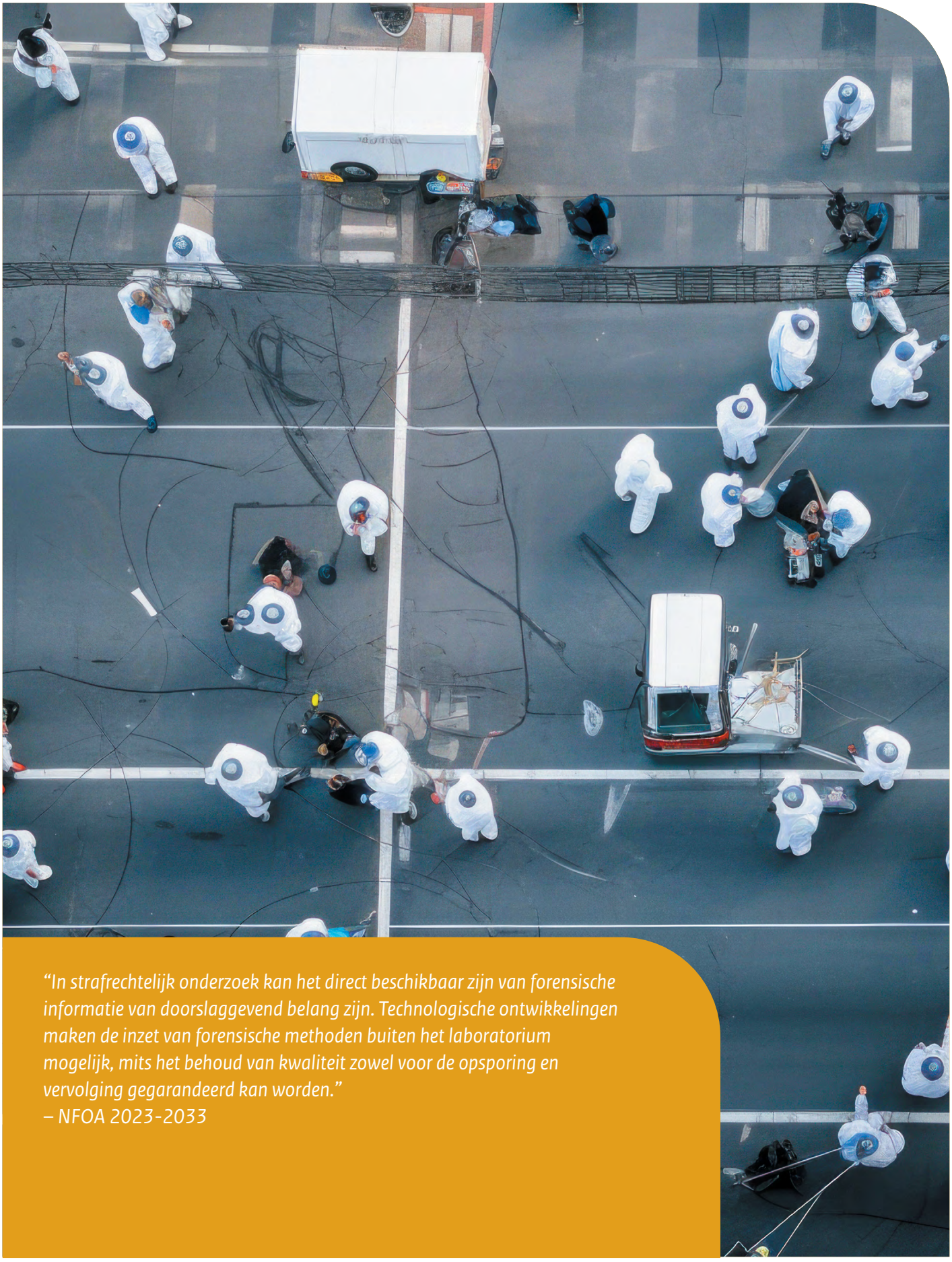
3D-reconstructie chips met SEM

Moderne apparaten, zoals smartphones, zijn steeds beter beveiligd. Om de gegevens inzichtelijk te maken moeten onderzoekers vaak niet alleen de beveiliging in de software omzeilen, maar ook de beveiliging in de chips zelf. Dat onderzoek begint met het maken van hoge resolutie foto's van de de verschillende lagen van de chip met behulp van een *Scanning Electron Microscope* (SEM). Deze microscoop maakt veel foto's die aan elkaar geplakt moeten worden. Voor het onderzoeken van de chip is het belangrijk dat de verschillende lagen ook netjes over elkaar heen worden gelegd. Dat doen bestaande oplossingen niet. Ook gaan bestaande oplossingen slecht om met grote stukken met repeterende patronen, en die komen op chips juist vaak voor. Daarom hebben we software ontwikkeld om computervision technieken slim toe te passen en daarmee de 3D-structuur van een chip te reconstrueren.

Objectivering morfologisch onderzoek

Is de verdachte in aanraking geweest met de sloot waarin het slachtoffer lag? Is deze schep gebruikt voor het graven van het graf? Door microscopische preparaten van bijvoorbeeld stuifmeel (uit grondsporen) en diatomeeën (uit water/modder sporen) automatisch te analyseren, kunnen deze en andere vragen sneller en beter worden beantwoord. De handmatige analyse van dergelijke microscopische preparaten levert nu al cruciale informatie op activiteitniveau in high-profile zaken, maar is ook een tijdrovende bezigheid die leunt op veel specialistische kennis. Om van een dag achter de microscoop voor één preparaat naar automatische analyse van acht preparaten buiten werktijd te gaan, is automatisch scannen en digitaal analyseren van preparaten samen met de Universiteit Utrecht verkend. Een preparaatscanner maakt daarvoor scans van stuifmeel-preparaten om daarna handmatig de stuifmeelkorrels te identificeren. Met deze data wordt een AI-model getraind om stuifmeel automatisch te herkennen. Het verkrijgen van digitale data, selectie van het juiste AI-model en de (wijze van) training vergt capaciteit voor data(base) opbouw, ontsluiting en beheer. Het ontwikkelde model kan daarna relatief snel worden uitgebreid naar andere typen preparaten, die kunnen verschillen in (3D) structuur, complexiteit en gelaagdheid/diepte.

Thema 3
Iedereen, overal, alles, altijd



"In strafrechtelijk onderzoek kan het direct beschikbaar zijn van forensische informatie van doorslaggevend belang zijn. Technologische ontwikkelingen maken de inzet van forensische methoden buiten het laboratorium mogelijk, mits het behoud van kwaliteit zowel voor de opsporing en vervolging gegarandeerd kan worden."
– NFOA 2023-2033



Iedereen, overal, alles altijd

ForMAT

Het project *Forensic Methylation Analysis Toolsets* (ForMAT) heeft als doel het verder brengen van het DNA-onderzoek in strafzaken, Disaster Victim Identification (DVI) en immigratie. Focus is op *DNA methylation* (DNAm) en het voorspellen van met name leeftijd van een persoon. Ook wordt gekeken naar de effecten van leefstijl en weefseltype. ForMAT ontwikkelt gespecialiseerde DNAm-kits op basis van geavanceerde sequencingtechnieken. Voor het ontwikkelen van bioinformatica-toepassingen worden statistische modellen gebruikt (er wordt geen gebruik gemaakt van AI of LLM's). Met de uitkomsten wordt gepoogd verbeterde leeftijdsvoorspelling te bereiken in straf- en DVI-zaken en een stralingsvrij alternatief te bieden voor leeftijd inschatting in immigratiezaken. Dit gebeurt in samenwerking met 11 partners bestaande uit universiteiten, politieorganisaties en forensische instituten uit verschillende Europese landen: Duitsland, Oostenrijk, Polen, Portugal, Spanje, het Verenigd Koninkrijk en Zweden.

Automatische tabletherkenning

Dit project ontwikkelt een tool om intelligence en inzicht te creëren in de opkomst, het uittreden en de geografische verdeling van MDMA-tabletten in Nederland. Hierdoor kunnen mogelijk relaties tussen zaken sneller zichtbaar worden. Dit wordt gedaan met geautomatiseerde vergelijkingen van foto's van MDMA-tabletten met beeldherkenning-AI. Wanneer alle forensische opsporingseenheden aangesloten zijn op de tabletherkenningstool en zij van ieder tablet een foto kunnen uploaden, zal er een landelijk beeld ontstaan. In de NFiDENT-software plukken deze eenheden hier gelijk de vruchten van, omdat bij iedere toevoeging er gelijk een vergelijking met de bestaande database wordt gemaakt en er teruggekoppeld wordt of de tablet eerder is gezien, hoe vaak en bij welke eenheid.

Explosieven in woonwijken - Verborgen verbanden onthullen

Voor strafrechtelijke onderzoeken waarbij sprake is van explosieven op basis van flitspoeder wordt in dit project een gevalideerde onderzoeksmethode ontwikkeld voor de bronvergelijking van flitspoeders. Met deze onderzoeksmethode wordt het mogelijk om verbanden binnen en tussen delicten zichtbaar te maken en om de sterkte van deze verbanden te bepalen. De basis van de methode zijn Likelihood Ratio-modellen die getraind zijn op verschillende kenmerken van flitspoeders. Tijdens de ontwikkeling wordt gekeken naar zowel de morfologische deeltjeskenmerken als de chemische samenstelling van flitspoeders die worden aangetroffen in Nederland.



Thema 4
AI & data science in de forensische praktijk



"De combinatie van kunstmatige intelligentie (AI) en grootschalige data biedt enorme mogelijkheden voor het forensisch werkveld. Een transparante aanpak is essentieel om de volle forensische potentie van de data science revolutie te benutten. AI methodes moeten volledig transparant en uitlegbaar zijn voor de samenleving en juridische professionals."
– NFOA 2023-2033



AI & data science in de forensische praktijk

Detector

Het project *Deepfake Evidence Technology for Forensic Content, Oversight and Research* (DETECTOR), zal bijdragen aan het ontwikkelen en verbeteren van forensisch onderzoek naar Deepfakes om zo de grote ondermijnende gevolgen hiervan te minimaliseren. Voor het NFI zal de nadruk liggen op het waarborgen van de bewijskracht van digitaal bewijs in de strafrechtketen. Dit project wordt uitgevoerd in samenwerking met collega's uit Duitsland, Finland, Griekenland, Italië, Polen, Portugal, Spanje en het Verenigd Koninkrijk.

MPS vanuit niet-humane matrices

De rol die niet-humane biologische sporen (NHBS) in zaakonderzoek hebben kan verschillend zijn, maar het DNA dat in cellen besloten ligt is een universele bron waarmee maximale informatie wordt verkregen uit deze sporen. Er zijn veel uiteenlopende DNA-technieken beschikbaar voor onderzoek aan NHBS. Nieuwe technologie die tot een grote verandering in het onderzoek van deze sporen leidt, is *massive parallel sequencing* (MPS). Deze technologie maakt het mogelijk om DNA-moleculen zowel zonder vermeerdering als na (a)specifieke vermeerdering direct uit te lezen. Mengsels van schimmels, bacteriën, virussen, planten en dieren zoals die voorkomen in forensische bemonsteringen en in grond, water en lucht, kunnen op basis van het aanwezige DNA in kaart worden gebracht tot een eDNA-profiel. Hiermee kunnen vragen als *'wat is het'* en *'waar komt het vandaan'* vaker, sneller en met hogere bewijskracht worden beantwoord.

De eerste toepassingen van MPS op het gebied van microbioom onderzoek worden al gevalideerd voor gebruik in zaakonderzoek. Om de techniek breder toe te passen en verder te implementeren wordt gewerkt aan doorontwikkeling van bio-informatica, data-engineering en datamanagement.

Agnostische deepfake-detectie

In het project *Agnostic Deepfake Detection* wordt vanuit Ondermijning gewerkt aan het ontwikkelen van detectie van AI gegenereerde afbeeldingen, zowel voor afbeeldingen gemaakt door gevestigde als nieuwe generatieve AI-modellen. Het doel is om een LLM te leren om AI-manipulatie te herkennen los van het specifieke model dat is gebruikt om de afbeelding te genereren of te bewerken. de generator die onder de afbeelding zit. Tijdens dit project wordt niet alleen aandacht besteed aan het trainen van het model, maar wordt ook gekeken naar de efficiëntie en schaalbaarheid. Dit doen we door de wisselwerking tussen nauwkeurigheid van detectie en rekencapaciteit in de gaten te houden.

Hansken RAG

Van september tot en met december 2025 werkten we vanuit het NFI samen met het Amsterdamse bedrijf Zeta Alpha aan het project Hansken RAG (Retrieval Augmented Generation). Dit deelproject van de Hansken Copilot onderzoekt hoe Hansken niet alleen een lijst met zoekresultaten kan tonen, maar ook direct een onderbouwd antwoord kan geven op de vraag van de onderzoeker, gebaseerd op relevante digitale sporen uit de zaak. RAG combineert het ophalen van passende informatie uit Hansken met de redeneer- en taalvaardigheid van een language model. Daardoor kan niet alleen gezocht, maar ook samengevat, uitgelegd en beantwoord worden. De nadruk ligt op transparantie en herleidbaarheid: antwoorden moeten terug te voeren zijn op de onderliggende bronnen. Het project verkent daarmee een volgende stap in AI-ondersteund digitaal forensisch onderzoek.



AI & data science in de forensische praktijk

Wachtwoordanalyse op schaal

Om toegang tot versleutelde data te krijgen in opsporingsonderzoeken is het vaak van belang om een wachtwoord te achterhalen. Een ingesteld wachtwoord laat sporen achter op een datadrager. Naast klassieke toepassingen om wachtwoordinformatie te vinden is er nu ook een AI-model ontwikkeld. Dit model sorteert ongestructureerde data zodat de tekst die het meest op een wachtwoord lijkt bovenaan in de lijst terecht komt. Met deze informatie kunnen onderzoekers effectievere woordenlijsten maken, die het kraken van een wachtwoord kansrijker maken.

AI voor DNA-analyse

DNA-profielanalyse is vaak arbeidsintensief. Hoewel geautomatiseerde verwerking al wordt toegepast, bijvoorbeeld binnen de *Snelle DNA Identificatie Lijn* (SIDL), voldoen sommige profielen nog niet aan de kwaliteitscriteria en blijft handmatige beoordeling nodig. Met behulp van *Deep Learning* worden modellen ontwikkeld die deze analyse efficiënter en consistenter uitvoeren. Betrouwbare, goed geannoteerde datasets zijn daarbij cruciaal. Daarom worden hulpmiddelen ontwikkeld om deze datasets efficiënt en zorgvuldig te creëren. Dit versnelt validatie en innovatie binnen forensisch DNA-onderzoek. Geavanceerde modellen maken het in de toekomst mogelijk meer profielen automatisch te verwerken en gevoeliger analyses uit te voeren, waardoor meer informatie uit DNA-sporen wordt gehaald en meer DNA-donoren kunnen worden opgespoord.



Thema 5
Ketenbreed meer met forensische informatie



"Grootschalig forensisch onderzoek biedt waardevolle informatie over criminaliteit en de effectiviteit van het strafrechtelijk onderzoek. Door te focussen op zaaksoverstijgende verbanden kunnen forensische gegevens een bron van informatie worden om criminaliteit te begrijpen en te bestrijden, en om forensische onderzoeksmethoden op ketenniveau te optimaliseren."
– NFOA 2023-2033



Ketenbreed meer met forensische informatie

Forensische analyse van geldstromen in ondermijnende criminaliteit

In de strijd tegen witwaspraktijken worden een nieuwe AI-tool en visuele analytische tools ontwikkeld. Deze AI-tool *AMLVis*, is een interactief hulpmiddel om de onderzoekers/rechercheurs/forensen te ondersteunen bij het examineren van diverse zaak data. Hiermee kunnen relaties in witwasnetwerken worden blootgelegd, mede dankzij het detecteren van abnormaliteiten in transactiedata. Met LLM's wordt de tool getraind en kunnen de onderzoekers focussen op het uitpluizen van de aangegeven relaties. Binnen het project wordt gefocust op de bruikbaarheid van de interactieve tool.

DNA Beslisondersteuning

De interpretatie van DNA-resultaten kan complex en tijdrovend zijn en leiden tot verschillen tussen en binnen laboratoria. Sporen die onvoldoende informatief blijken, kunnen worden onderworpen aan aanvullend onderzoek, waarna het verkregen DNA-profiel mogelijk wél geschikt is voor bijvoorbeeld opname in de DNA-databank voor Strafzaken. Bestaande combinaties van onderzoek kosten capaciteit en schaars DNA-extract, maar het niet toepassen van aanvullend onderzoek kan leiden tot gemiste opsporingsinformatie of verminderde bewijskracht. Wanneer is aanvullend onderzoek dan kansrijk? Dit onderzoek richt zich op datagedreven besluitvorming door te leren van historische (NFI) zaakdata en het ontwikkelen van systemen die keuzes ondersteunen, opleiding versterken en ketenpartners beter informeren over kansen en verwachtingen.

From anonymous to identified

Het ontbreken van de identiteit van een onbekende dode vormt potentieel een groot probleem voor het overlijdensonderzoek en de opsporing. Dit project beoogt met het gebruik van nieuwe AI-technieken te komen tot sturingsinformatie met betrekking tot de identiteit van een onbekende dode en daardoor de mogelijkheid te bieden tot nader (forensisch) onderzoek. Er zal een LLM/AI model worden getraind op basis van foto's gemaakt van overleden personen. Het model zal aan de hand van die foto's leren om tekeningen van het gezicht te maken. Zo'n referentie afbeelding/schets kan herkend worden en daarmee bijdragen aan identificatie.



Toekomstperspectief

De diversiteit aan projecten in dit portfolio is groot. Zowel in onderwerpen als in inzet. Van stageprojecten tot meerjarige internationale samenwerkingen, ze dragen allemaal bij aan één doel: beter forensisch onderzoek en daarmee een veilige en rechtvaardige samenleving. Hierbij zijn we bij het NFI gericht op de waarheid en gedreven door wetenschap. Dit moet tot verantwoorde, transparante en toetsbare toepassingen leiden.
De verantwoorde inzet van LLM's en andere mogelijkheden van data science is, naast een NFI-brede opgave, ook een opgave voor de hele keten. Daarom trekken we bij veel van onze innovaties samen op met de andere partijen in de strafrechtketen en stellen we onszelf steeds opnieuw de vraag: Leveren onze innovaties ook voor hen meerwaarde? Door de strategische prioriteiten van de forensische keten op te nemen in onze eigen Strategische Kennis- en Innovatieagenda (SKIA) en door behoeftegestuurde invulling van ons innovatieportfolio willen we zo veel mogelijk impact maken. Nu en in de toekomst.

Bottenbot

Het ontwikkelen van een *vision model* om voor het team Opsporen, Bergen, en Identificeren (OBI) een eerste triage te kunnen doen bij botten – zijn dit dierlijke of menselijke resten? Bij het vinden van botten bepaalt de politie of er een strafrechtelijk onderzoek moet volgen, waarbij de eerste vraag is of het een dierlijk of menselijk bot betreft. Dit kan nu veel tijd kosten van experts bij Politie en het NFI. Een model dat meekijkt met forensische experts van de Politie wanneer een foto wordt ingezonden kan hierbij het proces versnellen. De focus van dit project is om vooral de dierlijke botten te kunnen afvangen om onnodig werk te voorkomen. Wanneer er twijfel bestaat moet beslist nog teruggevallen worden op experts die het bot daadwerkelijk ook gaan bekijken.

Bevraagbare rapporten

Het project 'bevraagbare rapporten' heeft als doel door middel van LLM's rapporten van forensische geneeskunde en pathologie bevraagbaar te kunnen maken voor R&D en/of zaakoverstijgende vragen. Automatische pseudonimisering van de rapporten is hierbij de eerste grote stap zodat de rapporten anoniem en onherleidbaar zijn voor de onderzoeker. Daarnaast wordt uitgezocht of de rapporten veilig verwerkt kunnen worden op een vertrouwde Cloud omgeving. In de Cloud zou er met krachtigere LLM-modellen gewerkt kunnen worden, waardoor de antwoorden die de LLM geeft kwalitatief beter zijn.

Bijlage I Projectdetails

Project	Financiering	Discipline NFI	Samenwerkingspartners	Looptijd
SGPDO	Eigen en keten middelen	Opsporing aan de voor-kant van het proces	OM, Politie, Microsoft	Meerjarig programma
Speech-to-text	Eigen middelen	AI en data science	Intern	1 jaar
Robotiseer referentie-experiment	Programma	AI en data science	Intern	3 maanden
3D reconstructie chips met SEM	Eigen middelen	Betrouwbaar en snel onderzoek	Intern	1 jaar
Fractuurstudie	Eigen middelen	Van dataverzameling naar data gebruik	UvA, WKZ Utracht, Erasmus MC	1 jaar
Verklaarbare detectie van synthetisch spraakmateriaal	Eigen middelen	AI en data science	OM	8 maanden
ForMAT	Horizon Europe	Opsporing, aan de voorkant van het proces	Duitsland, Spanje, Oostenrijk, Portugal, Polen, Zweden	4 jaar
Automatische tabletherkenning	Eigen middelen	AI en data science	Politie	3 jaar
Detector	Horzion Europe	AI en data science	VK, Spanje, Polen, Finland, Italië, Portugal, Griekenland, Duitsland	3 jaar
Explosieven in woonwijken - Verborgen verbanden onthullen	Eigen middelen	Van dataverzameling naar datagebruik	Intern	4 jaar
Objectivering morfologisch onderzoek	Eigen middelen	Betrouwbaar en snel onderzoek	Universiteit Utrecht	Doorlopend
MPS vanuit niet-humane matrices	Eigen middelen	Betrouwbaar en snel onderzoek	Intern	Doorlopend
AI voor DNA-analyse	Eigen middelen	AI en datascience	Intern	1 jaar
Bevraagbare rapporten	Eigen middelen	AI en data science	Intern	3 jaar
Agnostische Deepfake-Detectie	Programma	AI en data science	UvA, AI4Forensics	2 jaar
From anonymous to identified	Externe middelen	Betrouwbaar en snel onderzoek	ETVR, ATOE, LOEP, CvB	1 jaar
Forensische analyse van geldstromen in ondermijnende criminaliteit	Programma	AI en datascience	Politie, AMLC, Deloitte, UvA	4 jaar
Wachtwoordanalyse op Schaal	Programma	AI en datascience	Diverse partners veiligheids-domein	8 maanden
DNA Beslisondersteuning	Eigen middelen	AI en datascience	OM, Politie, ZM	3 jaar
Hansken RAG	Hansken	AI en datascience	Hansken community partners	Meerjarig programma
Hansken Copilot	Hansken	AI en datascience	Hansken community partners	Meerjarig programma
Bottenbot	Eigen middelen	AI en data science	OM, Politie	2 jaar

Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid

Postbus 24044 | 2490 Den Haag
Telefoon (070) 888 66 66

https://www.forensischinstituut.nl

© Rijksoverheid, versie juli 2026