



**kwiteit  
forensische  
zorg**

2022

# AI binnen 3RO

Toepassing en toekomst

Dr. Iris van Sintemaartensdijk  
Marleen Haandrikman, MSc  
Dr. Hanneke Kip  
Dr. Caroline Fischer  
Prof. dr. René Torenvlied

Datum van publicatie: september 2022

## **Colofon**

Deze publicatie/dit project is tot stand gekomen met financiële ondersteuning van het Programma KFZ. Het Programma KFZ heeft daarnaast zorg gedragen voor de inhoudelijke projectbegeleiding waarbij is gestuurd op kwaliteit, voorgangsbewaking en budgetbeheer. Het Expertisecentrum Forensische Psychiatrie voert namens het programma KFZ het projectmanagement.

© KFZ: Het intellectuele eigendom ligt bij de auteur(s). De auteur(s) stemt er mee in dat deze uitgave onvoorwaardelijk en zonder kosten gebruikt kan worden door alle instellingen binnen het forensische zorgveld

## Inhoudsopgave

1. Samenvatting .....	4
1.1 Nederlands .....	4
1.2 Engels .....	5
2. Inleiding .....	6
2.1 AI binnen 3RO .....	6
2.2 Doelstellingen en onderzoeksvragen .....	6
2.3 Opzet van het rapport .....	7
3. Methoden .....	8
3.1 Opzet scoping review .....	8
3.2 Opzet interviewstudie met experts .....	9
3.3 Opzet interviewstudie met reclasseringswerkers .....	9
4. Resultaten .....	11
4.1 Resultaten scoping review .....	11
4.1.1 Definities AI .....	11
4.1.2 AI in reclasseringswerk .....	12
4.1.3 AI binnen justitie en GGZ .....	14
4.1.4 Aspecten in acht te nemen bij implementatie van AI .....	15
4.1.5 Conclusie scoping review .....	17
4.2 Resultaten interviews met experts .....	17
4.2.1 Definities van AI .....	18
4.2.2 Toepassingen AI binnen reclassering en het justitiële veld .....	18
4.2.3 Morele, ethische en juridische aspecten aangaande toepassing AI binnen 3RO .....	19
4.2.4 Stappen nodig om AI binnen 3RO succesvol in te kunnen zetten .....	20
4.2.5 Conclusie expert interviews .....	21
4.3 Resultaten interviews met reclasseringswerkers .....	22
4.3.1 Kennis van AI .....	22
4.3.2 Huidig gebruik van AI binnen 3RO .....	22
4.3.3 Mogelijke toepassingen AI binnen 3RO .....	24
4.3.4 Stappen nodig om AI binnen 3RO succesvol te implementeren .....	28
4.3.5 Conclusie interviews met reclasseringswerkers .....	29
5. Discussie .....	31
5.1 Huidige toepassing AI binnen justitiële veld en 3RO (onderzoeksvraag 1 en 3) ....	31
5.2 AI door de ogen van de reclasseringswerker (onderzoeksvraag 2) .....	31
5.3 De toekomst in toepassing van AI binnen 3RO (onderzoeksvraag 4 tot en met 8)	32
5.4 Concrete aanbevelingen .....	32
5.4.1 Voorbij AI als risico-taxatie .....	32

5.4.2 Kijk naar de ontwikkelingen binnen (explainable) AI .....	33
5.4.3 Verlies de menselijke maat niet .....	33
5.5 Conclusie.....	33
6. Referenties .....	34
Appendix A: Scoping review .....	37
Appendix B: Protocol interview AI-experts .....	50
Appendix C: Protocol interview reclasseringswerkers .....	52

# 1. Samenvatting

## 1.1 Nederlands

Artificial intelligence (AI) speelt een steeds grotere rol in ons leven en wordt langzamerhand in meer branches en organisaties ingezet. Ook binnen het justitiële veld worden de mogelijkheden onderzocht om AI in te kunnen zetten. Een belangrijke verkenning hierbinnen is de mogelijke toepassing van AI binnen 3RO. Het doel van dit rapport is om voor het reclasseringswerk een ruimer begrip te krijgen over wat AI is, hoe AI ingezet kan worden binnen dit reclasseringswerk, wat de voor- en nadelen van deze toepassingen zijn en hoe reclasseringswerkers met zulke AI-programma's kunnen werken.

Om dit te onderzoeken werd een scoping review uitgevoerd en interviews gehouden met AI-experts en reclasseringswerkers. Hieruit blijkt dat AI internationaal nog weinig toegepast wordt binnen de reclassering, behalve in risico-taxatie instrumenten. AI gebruiken om informatie te ordenen en aan te bieden aan reclasseringswerkers is een toepassing die op ondersteuning kan rekenen van de reclasseringswerkers. Wanneer AI-systemen goed functioneren kunnen deze potentieel werk uit handen nemen. Echter, nadelen aan de inzet van AI zijn mogelijke bevooroordeeldheid in de data, niet weten hoe het algoritme precies beslissingen maakt en de acceptatie van AI-systemen door gebruikers. Het is duidelijk dat AI-systemen constant moeten worden gemonitord om dit tegen te gaan. Binnen 3RO moet AI vooral de reclasseringswerkers kunnen ontzorgen en ontlasten om succesvol te worden ingezet. De menselijke maat behouden in het reclasseringswerk is hierbij essentieel.

Het is duidelijk dat AI binnen het justitiële veld in opkomst is. Steeds meer AI-programma's worden ontwikkeld en in gebruik genomen. Voor 3RO lijken er kansen te liggen in het gebruik van AI om reclasseringswerkers te ontlasten in de informatie die zij moeten verwerken. Duidelijk is dat er nog veel haken en ogen zitten aan de inzet van AI-systemen. Concreet bevelen wij aan dat wanneer men verder wil met AI binnen 3RO er (1) voorbij AI als risico-taxatie moet worden gekeken, (2) gekeken moet worden naar de continue ontwikkelingen van AI-systemen en wat er mogelijk is, en (3) de menselijke maat in het oog houden zodat AI-systemen nooit het eindoordeel geven. We zitten voor het toepassen van AI binnen het justitiële veld – en binnen 3RO in het bijzonder – nog in de beginfase. We zullen kritisch moeten blijven op AI-systemen en de inzet van deze systemen continue moeten monitoren met het oog voor de menselijke maat.

## 1.2 Engels

Artificial intelligence (AI) has started to play a bigger role in our lives and is slowly more deployed in several branches and organisations. Within the judicial field, the opportunities to apply AI are being studied. An important exploration here is the possible application of AI within the Dutch probation services (3RO). The goal of this report is to provide the probation services with a better understanding of what AI is, how AI can be deployed within probation services, the pros and cons of the potential applications, and how probation service workers will work with such AI programmes.

To study this, we employed a scoping review and interviewed both AI experts and probation workers. We found that AI is scarcely applied internationally within the probation setting, apart from risk assessment instruments. Using AI to organise information is an application probation workers would find valuable. When these AI systems function properly, they can potentially ease the workload of the probation workers. However, downsides of AI systems are potential prejudice in data, not knowing how the algorithm makes a decision and accepting the decisions of AI-systems by the users (both probation workers and convicts). It is clear that these AI systems need to be constantly monitored to prevent this. Within the probation AI needs to be able to decrease the workload of the probation workers in order to be implemented successfully. Keeping the human factor front and central in the probation is essential in implementing AI systems.

It is clear that the application of AI systems is rising within the judicial field. There seem to be possibilities for the probation in using AI to unburden employees in processing information. It is clear that there are still a lot of issues however the with AI systems. We advise the following steps to be taken before implementing AI within the probation: (1) go beyond AI as risk taxation instrument, (2) keeping track of the developments of AI systems (both positive and negative) and (3) to put the human factor front and central and make sure that AI systems are never taking the final decisions. We are still in the beginning phase of applying AI within the judicial field, and especially within the probation. We need to stay critical of AI systems and need to continuously monitor these systems with the focus on maintaining the human factor.

## 2. Inleiding

In dit hoofdstuk wordt AI binnen 3RO geïntroduceerd en worden de onderzoeksvragen toegelicht die centraal zullen staan in dit onderzoek.

### 2.1 AI binnen 3RO

Alhoewel velen van ons niet precies weten wat artificial intelligence (AI) precies inhoudt, is het langzamerhand al een onderdeel van ons dagelijks leven. AI bepaalt bijvoorbeeld welke e-mails er door de spamfilter komen, hoe je via een app moet rijden naar een bepaalde bestemming of het suggereert producten die je kan bestellen in webshops. Er is tot op heden nog geen vaststaande definitie van AI, maar het wordt algemeen beschouwd als een set van technologieën die zorgen voor het verzamelen, analyseren en presenteren van een grote hoeveelheid data op een snelle manier (Stanford University, 2016). AI kan levens makkelijker maken door snel(ler) beslissingen te kunnen nemen waar de mens veel langer de tijd voor nodig heeft – en mogelijk niet alleen efficiënter maar zelfs ook accurater (Rowland et al., 2015). AI blijkt echter een “black box” te zijn (Deeks, 2018). Er komt informatie in het algoritme en er volgt een beslissing, maar hoe de AI beslissingen neemt blijft onduidelijk.

AI wordt langzamerhand in meer branches en organisaties ingezet. Ook binnen het justitiële veld worden de mogelijkheden onderzocht om AI in te kunnen zetten. Een belangrijke verkenning hierbinnen is de mogelijke toepassing van AI binnen reclasseringswerk. Als AI kan helpen betere beslissingen te nemen en efficiënter te werken, is het van belang om na te denken of dit binnen het reclasseringswerk ook van toepassing zou kunnen zijn. Het reclasseringswerk in Nederland wordt uitgevoerd door drie reclasseringsorganisaties (3RO): (1) Reclassering Nederland, (2) Stichting Verslavingsreclassering GGZ, en (3) Stichting Leger des Heils Jeugdbescherming & Reclassering<sup>1</sup>. De drie organisaties verschillen van elkaar wat betreft geschiedenis, visie, omvang, structuur en manier van werken. In de praktijk werken ze nauw met elkaar samen om bij te dragen aan de veiligheid van de samenleving door het voorkomen van criminaliteit en het terugdringen van recidive. Binnen 3RO wordt er een perspectief gehanteerd waarin 4 “producten” samenkomen: (1) Advies, (2) Toezicht, (3) Werkstraffen, en (4) Justitiële interventies. Deze moeten worden samen gezien als één traject die de re-integratie van een cliënt bevordert. Dit samenhangende traject wordt langzamerhand meer en meer geïntegreerd in de praktijk. Het doel voor 3RO is om expert te worden omtrent de afbouw van criminele carrières.

Het gebruik van goed werkende AI zou binnen 3RO een ondersteuning kunnen zijn voor de reclasseringswerkers. Echter, het is niet duidelijk wat de toevoeging van AI-programma’s zou kunnen betekenen voor de reclasseringswerkers en welke vormen van AI er tot de beschikking staan. Daarnaast is de vraag hoe de reclasseringswerkers voorspellingen en beslissingen van AI moet wegen tegenover het eigen oordeel. AI lijkt veel mogelijkheden te hebben, maar het is een nog veel onontgonnen gebied.

### 2.2 Doelstellingen en onderzoeksvragen

Het doel van dit rapport is om voor het reclasseringswerk beter te begrijpen wat AI is, hoe AI ingezet kan worden binnen dit reclasseringswerk, wat de voor- en nadelen van deze toepassingen zijn en hoe reclasseringswerkers met zulke AI-programma’s kunnen werken. Op basis hiervan kunnen concrete adviezen gegeven worden over hoe en in welke mate AI binnen 3RO kan worden ingezet.

De volgende onderzoeksvragen staan hierbij centraal:

1. Wat is de huidige stand van zaken in AI in een justitiële en GGZ-context, en wat is de meest passende definitie van AI?
2. Hoe wordt AI gezien door medewerkers binnen de 3RO?
3. Waar en hoe wordt AI al binnen de context van het reclasseringswerk toegepast?
4. Wat zijn mogelijke toepassingsmogelijkheden van AI binnen het reclasseringswerk?
5. Wat zijn mogelijke voor- en nadelen van de verschillende toepassingsmogelijkheden van AI in het reclasseringswerk?
6. Wat zijn morele, ethische en juridische aspecten die een rol spelen in de beslissing om AI in het reclasseringswerk toe te passen?
7. Welke vervolgstappen zijn er nodig in onderzoek naar AI binnen het reclasseringswerk?

---

<sup>1</sup> Dienst Justitiële Inrichtingen (2022). *3 Reclasseringorganisaties (3RO)*. <https://www.forensischezorg.nl/introductie/keten-forensische-zorg/3-reclasseringsorganisaties-3ro>

8. Welke vervolgstappen zijn er nodig in de praktijk om AI goed toe te passen binnen het reclasseringswerk?

### 2.3 Opzet van het rapport

Om de onderzoeksvragen te beantwoorden en de doelstelling van het rapport te behalen is er literatuuronderzoek uitgevoerd en zijn zowel AI-experts als reclasseringswerkers geïnterviewd. Eerst beschrijven wij de methodologie in het onderdeel methoden. Vervolgens beschrijven wij de resultaten voortkomend uit de scoping review, de interviews met AI-experts en de interviews met reclasseringswerkers in het hoofdstuk resultaten. In het hoofdstuk conclusie worden conclusies getrokken over de staat van AI binnen 3RO en waar de toepassing en toekomst van deze programma's ligt. Tot slot worden er concrete adviezen gegeven over de inzet van AI binnen 3RO.



### 3. Methoden

Om de onderzoeksvragen te beantwoorden en de doelstelling van het rapport te behalen, is er literatuuronderzoek uitgevoerd en zijn zowel AI-experts als reclasseringswerkers geïnterviewd. De resultaten zullen dus drie onderdelen omvatten. In tabel 1 is te zien welke onderzoeksvraag in welk onderdeel wordt behandeld.

**Tabel 1.**

*Onderzoeksvragen behandeld per onderdeel van de resultaten*

<b>Onderzoeksvragen</b>	<i>Scoping review</i>	<i>Interviews met AI-experts</i>	<i>Interviews met reclasseringswerkers</i>
1. Wat is de huidige stand van zaken in AI in een justitiële en GGZ-context, en wat is de meest passende definitie van AI?	✓	✓	
2. Hoe wordt AI gezien door medewerkers binnen de 3RO?			✓
3. Waar en hoe wordt AI al binnen de context van het reclasseringswerk toegepast?	✓	✓	✓
4. Wat zijn mogelijke toepassingsmogelijkheden van AI binnen het reclasseringswerk?	✓	✓	✓
5. Wat zijn mogelijke voor- en nadelen van de verschillende toepassingsmogelijkheden van AI in het reclasseringswerk?	✓	✓	✓
6. Wat zijn morele, ethische en juridische aspecten die een rol spelen in de beslissing om AI in het reclasseringswerk toe te passen?	✓	✓	✓
7. Welke vervolgstappen zijn er nodig in onderzoek naar AI binnen het reclasseringswerk?	✓	✓	✓
8. Welke vervolgstappen zijn er nodig in de praktijk om AI goed toe te passen binnen het reclasseringswerk?	✓	✓	✓

#### 3.1 Opzet scoping review

Om de huidige stand van zaken omtrent het gebruik van AI en de toepassing in het reclasseringswerk in kaart te brengen, werd een verkennend systematisch literatuuronderzoek (scoping review) uitgevoerd van artikelen binnen de nationale en internationale literatuur.

Bij een scoping review wordt de onderzoeksstrategie, inclusie- en exclusiecriteria en de wijze van coderen van te voren gespecificeerd (Watson et al., 2015). De inclusiecriteria en reikwijdte van de review worden daarna verder aangescherpt tijdens het dataverzamelingsproces. Hierbij worden meerdere typen papers geïnccludeerd, waaronder ook "grijze" literatuur (literatuur gepubliceerd door o.a. overheidsinstanties maar niet peer-reviewed)". Het doel van een scoping review is om een overzicht te geven van de stand van zaken binnen een veld. Het verschil met een systematic review is dat er bij een scoping review geen kwaliteitscheck is van de papers, omdat het doel is een overzicht te geven van de huidige stand van zaken.

Voor de scoping review werd allereerst gekeken naar literatuur over AI binnen het reclasseringswerk. Daarnaast werd er gezocht naar literatuur over AI binnen de justitiële context in het algemeen en binnen de GGZ. Hiervoor werd een search string opgesteld (zie Appendix A, tabel 2), met kernwoorden zoals "probation", "artificial intelligence" en "justice". Deze search string werd verfijnd om irrelevante papers uit te sluiten. Om zo veel mogelijk disciplines mee te nemen in de zoektocht werden er vijf academische databases doorzocht: Scopus, Web of Science, PsycInfo, PubMed en

ACM. Grijze literatuur werd gezocht door websites van stakeholders te bezoeken en via gesprekken met andere onderzoekers.

Om de *abstracts* van artikelen te screenen op relevantie voor de onderzoeksvragen werd de AI-tool ASReview gebruikt. Met deze tool is het mogelijk om artikelen snel te screenen op relevantie, door alle papers in het programma te zetten en daarna het programma te leren welke papers relevant en irrelevant zijn. Hiervoor werden 10 relevante papers ingevoerd met als onderwerp de toepassing van AI binnen het justitiële veld wat was geselecteerd door de onderzoekers. Daarna werden er 5 irrelevante papers willekeurig geselecteerd uit de databases van artikelen. Nadat ASReview deze papers had gescreend, bood het vervolgens één-voor-één de overgebleven papers aan. Elk paper werd individueel als relevant of irrelevant beoordeeld. Door elk beoordeelde paper leert de tool welke papers zeer waarschijnlijk het meest relevant zijn, en biedt vervolgens deze weer in de meest optimale volgorde aan. Artikelen werden uitgesloten als deze: (1) zich niet richten op AI binnen de reclasseringspraktijk of (2) AI binnen de GGZ of (3) binnen justitie. Daarnaast werden papers ook uitgesloten als deze alleen over de technische kant van algoritmes en AI gingen in plaats van over de toepassing van dit algoritme. Na de abstract-screening van de artikelen werden de volledige papers gelezen, waarbij dezelfde exclusiecriteria werden toegepast.

### 3.2 Opzet interviewstudie met experts

Om te onderzoeken wat de stand van zaken is rondom de toepassingen van AI binnen de reclassering en het justitiële veld zijn er interviews afgenomen met drie experts.

De experts zijn prof. dr. Tibor Bosse (hoogleraar communicatiewetenschap en kunstmatige intelligentie aan de Radboud Universiteit Nijmegen), prof. dr. Peter-Paul Verbeek (hoogleraar filosofie van technologie aan de Universiteit Twente) en dr. Leendert van Maanen (universiteit hoofddocent cognitieve aspecten van artificial intelligence aan de Universiteit Utrecht). De vragen voor het interviewprotocol zijn opgesteld aan de hand van de scoping review. In bijlage C kunt u het interviewprotocol vinden.

Alle deelnemers kregen voor het interview een presentatie toegestuurd met daarin een korte uitleg over 3RO en wat er al aan AI gebruikt wordt. Deze presentatie was optioneel en hoefde niet doorgekeken worden. Alle deelnemers werden online geïnterviewd om reistijd te beperken. Nadat geverifieerd werd of deze presentatie was bekeken of niet, werd er gestart met het interview. Alle vragen werden tijdens het interview gesteld, maar de volgorde kon worden veranderd wanneer dit uitkwam (als er bijvoorbeeld een relevant punt werd aangehaald wat normaliter later besproken zou worden). Als laatste punt werd er altijd gevraagd naar input van de expert en of alles besproken was. De interviews duurden ongeveer 60 minuten.

De interviews zijn allen getranscribeerd via AmberScript en doorgelopen om zo tot een zorgvuldige transcriptie te komen. Hierna werden de interviews thematisch gecodeerd.

### 3.3 Opzet interviewstudie met reclasseringswerkers

Om te onderzoeken hoe reclasseringswerkers de huidige toepassing van AI binnen 3RO zien en waar zij kansen en moeilijkheden verwachten, werden er interviews afgenomen.

In totaal zijn er zes reclasseringswerkers binnen 3RO geïnterviewd. Deze interviews zijn anoniem afgenomen. De jaren ervaring binnen 3RO van de reclasseringswerkers liggen tussen de 12 en 35 jaar. Van hen zijn er vijf werkzaam binnen de Reclassering Nederland en één binnen de Stichting Verslavingsreclassering GGZ. Alhoewel niet iedere geïnterviewde medewerker op dit moment actief is als reclasseringswerker – en dus niet actief met cliënten werkt – zijn allen dat wel in het verleden geweest. De vragen voor het interviewprotocol zijn opgesteld aan de hand van de scoping review en de interviews met experts. In bijlage C kunt u het interviewprotocol vinden.

Alle deelnemers kregen voor het interview een presentatie toegestuurd met een korte uitleg over AI en wat er momenteel waarschijnlijk al aan AI binnen 3RO gebruikt wordt. Deze presentatie was optioneel en hoefde niet doorgekeken worden. Alle deelnemers werden online geïnterviewd om reistijd te beperken. Nadat geverifieerd werd of deze presentatie was bekeken of niet, werd er gestart met het interview. Alle vragen werden tijdens het interview gesteld, maar de volgorde kon worden veranderd wanneer dit uitkwam (als er bijvoorbeeld een relevant punt weer aangehaald wat normaliter later besproken zou worden). Als laatste punt werd altijd gevraagd naar input van de reclasseringswerker en of alles besproken was. De interviews duurden ongeveer 45 minuten.

De interviews zijn allen getranscribeerd via AmberScript en doorgelopen om zo tot een zorgvuldige transcriptie te komen. Hierna werden de interviews thematisch gecodeerd.

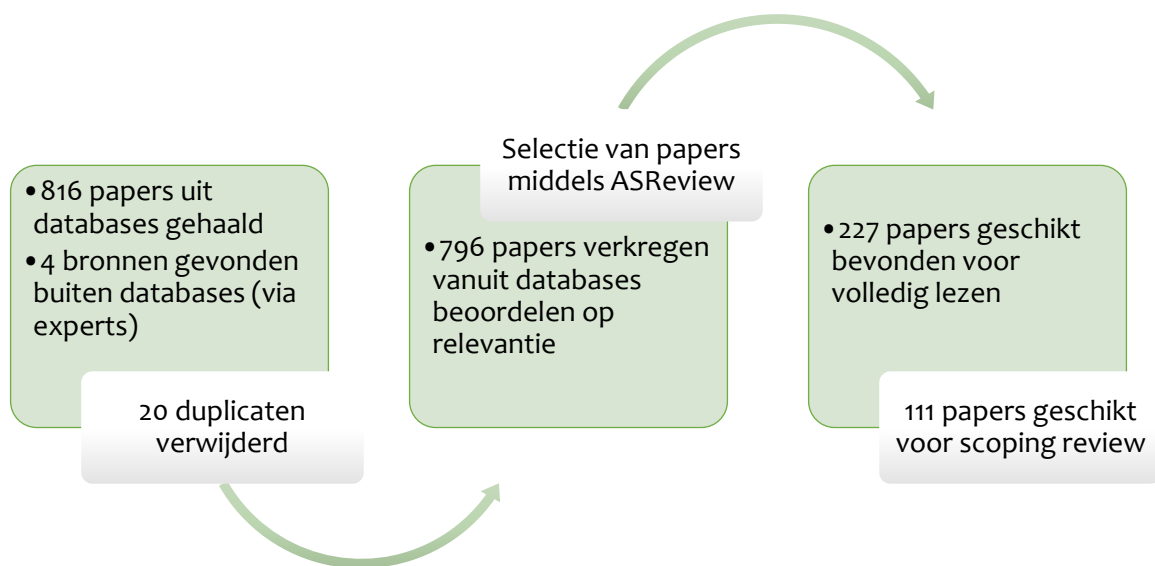
## 4. Resultaten

Hieronder beschrijven wij de resultaten van de drie onderdelen van dit project: de scoping review, de interviews met AI-experts en de interviews met reclasseringswerkers.

### 4.1 Resultaten scoping review

Vanuit de zoektocht werden 796 papers uit de databases gehaald. Figuur 1 vat samen hoeveel papers werden uitgesloten tijdens elk stadium van het onderzoek. Uiteindelijk bleven er 107 papers over, waarbij nog vier overige bronnen werden toegevoegd die aangeraden werden door andere onderzoekers. Dit betroffen twee academische papers en twee papers die uit de "grijze" literatuur komen. In tabel 3 (Appendix A) worden alle papers getoond.

Van deze papers hebben er 84 de toepassing van AI binnen het justitiële veld als onderwerp. Er zijn 18 papers die over de toepassing van AI binnen meerdere velden als onderwerp hebben, waarin het justitiële veld één van deze velden is. Verder hebben er vijf papers de toepassing van AI binnen reclassering als onderwerp en vier papers de toepassing van AI binnen de GGZ.



Figuur 1: Overzicht van selectieproces voor gebruik van papers in het scoping review

Wij bespreken de resultaten hieronder in het licht van de onderzoeksvragen. Allereerst bespreken we de meest passende definitie van AI (onderzoeksvraag 1), de toepassing van AI binnen de context van de reclasseringspraktijk (onderzoeksvraag 3 en 4), toepassing van AI binnen justitiële en GGZ-context (onderzoeksvraag 1), en aspecten in acht te nemen bij het implementeren van AI binnen de reclassering en het justitiële veld (onderzoeksvragen 5 en 6).

#### 4.1.1 Definities AI

Niet elk paper wat over AI gaat, haalt ook een definitie van AI aan. Van de 111 papers zijn er 24 met een concrete definitie van AI. Dit is opvallend, omdat consistentie in wat AI precies inhoudt wel een voorwaarde lijkt om te begrijpen welke programma's wel en niet geclassificeerd kunnen worden als AI. Tabel 4 (Appendix A) laat zien wat deze definities van AI zijn.

De kern van de definities zit in het uitvoeren van taken geïnspireerd door taken die mensen zouden kunnen uitvoeren, maar ook taken die buiten het bereik liggen van ons menselijk kunnen. Het Dutch Artificial Intelligence Manifesto geschreven door de Special Interest Group on Artificial Intelligence biedt een goed overzicht van de definitie van AI (2018): "Artificial Intelligence (AI) is the science and engineering that studies and creates intelligent systems".

Zij benoemen 7 kerngebieden binnen het AI-onderzoek.

- Autonome agents en robotica: computer systemen die in (fysieke of online) systemen werken om hun doelen te behalen;
- Computer visualisatie: een visuele representatie van de wereld maken via AI;
- Beslisprocessen: systemen die regel-gebaseerd zoeken en beslissingen maken;
- Informatie ophalen: AI-systemen die mensen met informatie verbinden;
- Kennis representatie en redenering: AI-systemen die informatie weergeven en verwerken om complexe taken op te lossen;
- Machine learning: AI systemen die niet alleen informatie verwerken maar ook leren van de taak die ze uitvoeren;
- Natural language processing: AI-systemen die informatie en kennis ophalen uit gesproken of geschreven woord.

Om de volgende stap in het gebruik en onderzoek van AI te kunnen maken, zijn volgens de Special Interest Group on Artificial Intelligence (2018) drie multidisciplinaire uitdagingen relevant:

1. Socially-Aware AI: dit zijn AI-systemen die ontwikkelt worden om proactief en op een persoonlijke manier samen te werken met mensen;
2. Explainable AI: dit zijn AI-systemen die beslissingen kunnen nemen die voorheen enkel alleen door mensen konden worden uitgevoerd, waarbij de beslissingen een belangrijke impact hebben op het leven van mensen. Van belang is hierbij dat het systeem uit kan leggen waarom het deze beslissing heeft gemaakt;
3. Responsible AI: dit zijn AI-systemen die taken die voorheen mensen alleen konden uitvoeren gaan automatiseren, alsmede ook het automatisch verwerken van een grote hoeveelheid data.

AI-systemen hebben momenteel nog niet de capaciteit om uit te leggen waarom beslissingen genomen worden (Special Interest Group on Artificial Intelligence, 2018). Mogelijk zou de acceptatie van beslissingen die gedeeltelijk genomen worden door AI hoger worden als het systeem uitlegt waarom deze beslissingen uiteindelijk tot stand kwamen.

De resultaten van de scoping review over de definitie van AI laat ons zien dat er geen algemeen geldende definitie gehanteerd wordt, maar dat de kern zit in het ontwerpen van intelligente systemen die net zo veel of nog meer kunnen dan mensen.

#### 4.1.2 AI in reclasseringswerk

Vanuit de scoping review komt naar voren dat er een klein aantal papers zijn die specifiek het gebruik van AI binnen de reclassering als onderwerp hebben. In totaal waren dit vijf papers. In tabel 5 (Appendix A) worden deze papers verder toegelicht.

Het eerste onderzoek is van Bicacki, Seker en Demirel (2020). Zij bestudeerden de spatio-temporele patronen van voorwaardelijk vrijgelatenen die onder toezicht gesteld waren en die een enkelband droegen in Istanbul, Turkije. De reclassering in Turkije is pas sinds 2005 actief (Erdem et al., 2019). Momenteel lijkt Turkije nog voor belangrijke uitdagingen te staan in het ontwikkelen van het reclasseringswerk, waaronder het opzetten van rehabilitatieprogramma's voor voormalig gedetineerden en het huidige gebrek aan training voor reclasseringswerkers in onder andere risicotaxatie. Reclasseringswerkers in Turkije geven aan dat zij hun twijfels hebben bij hun rol binnen het justitiële systeem, waarbij ze een gevoel van machteloosheid ervaren (zie Erdem en collega's (2019) voor een uitgebreidere rapportage over de staat van de Turkse reclassering). Bicacki, Seker en Demirel (2020) wilden graag weten wat de beweegpatronen van de enkelbanddragers waren en op welke locaties zij de meeste tijd doorbrachten. 77 enkelbanddragers werden per minuut gevolgd, waarna een AI-systeem berekende waar deze enkelbanddragers hun tijd doorbrachten en of dat geclassificeerd kon worden als aantrekkelijk gebied voor ze. De onderzoekers merkten op dat deze informatie waardevol kan zijn voor beleidsmakers wanneer deze uitgebreider wordt verzameld. Alhoewel dit een pilot onderzoek betreft, gaan de onderzoekers niet nader in op de voor- en nadelen van het gebruik van deze data. De Turkse reclassering zit in een andere ontwikkelingsfase dan de Nederlandse reclassering, wat de vertaling van deze resultaten naar de Nederlandse context lastiger maakt. Het laat wel een toepassing zien van AI die binnen de Nederlandse reclassering ook mogelijk is, maar in het onderzoek wordt niet aangehaald waarom deze toepassing relevant is. Het mogelijk inzetten van informatie van enkelbanden binnen de Nederlandse reclassering in AI-systemen wordt verder besproken in de interviews met de reclasseringswerkers verderop in dit hoofdstuk.

Berk (2017) onderzocht in de Verenigde Staten de toepassing van een AI-systeem om te bepalen wie er onder toezicht komt te staan of in de gevangenis moet blijven. Nadat het systeem werd geïntroduceerd, namen de beslissingen die besloten wie er onder toezicht kwam te staan af van 62%

naar 58%. Tussen het AI-systeem en de besliscommissie zat weinig verschil in de gemaakte beslissingen. Ook werd het systeem als ondersteuning gebruikt en werden deze AI-beslissingen in de commissie besproken. Vergelijkingen tussen de beslissingen van de commissie met en zonder AI-systeem kunnen dus moeilijk gemaakt worden. De Amerikaanse reclassering heeft een groot aantal mensen onder toezicht, in 2019 was dit 1 op de 59 volwassen Amerikanen. In het Amerikaanse systeem staan mensen onder toezicht die een werkstraf uitzitten in plaats van celstraf, alsmede ook gedetineerden die hun celstraf hebben uitgezeten en vrijgelaten worden op voorwaarde van een (lang) toezicht. De Amerikaanse reclassering is beter vergelijkbaar met de Nederlandse reclassering ten opzichte van de eerder genoemde Turkse reclassering. Er zijn vergelijkbare taken in beide landen voor de reclasseringswerker, en het doel is re-integratie in de maatschappij (zie (Oudekerk & Kaeble, 2021 voor meer informatie over de Amerikaanse reclassering). Een nadere uitleg over de toepassing van een AI risico-taxatiesysteem dat voornamelijk in de Verenigde Staten wordt gebruikt volgt hieronder onder punt 4.1.3.1 in de scoping review.

Bagaric en collega's (2019) vergeleken beslissingen – gemaakt door zowel mensen als AI-systemen – over wie onder toezicht kwamen te staan van de reclassering of wie nog in de gevangenis moest blijven zitten. In hun onderzoek hebben zij gekeken naar verschillende AI-systemen die gebruikt worden in verschillende staten in de Verenigde Staten. Zij vinden dat de menselijke beslissingen niet optimaal functioneren, en er is een grote kans dat de voormalig gedetineerden weer de fout ingaan. Het huidige proces is niet accuraat en zeer complex. Ze benoemen actuariële risico-taxatie als een accurater middel om de recidive-kans te berekenen vergeleken met een ongestructureerd oordeel van deze recidive-kans. Ze worden echter nog niet veelvuldig ingezet. In het artikel wordt echter niet nader ingegaan op één systeem, en een bespreking van elk systeem in detail (over o.a. de voor- en nadelen van deze systemen en de vergelijking tussen deze systemen) wordt niet gedaan. Ook dit onderzoek werd voornamelijk in de Verenigde staten uitgevoerd, en risico-taxatie binnen de Verenigde staten wordt verder besproken onder punt 4.1.3.1 in de scoping review.

Binnen 3RO is er al een programma waarin AI wordt toegepast: RISC, het risicotaxatie- en adviesinstrument. Risico-taxatie instrumenten geven een schatting van het risico – in dit geval recidive – op basis van ingevoerde kenmerken van de cliënt. RISC betreft een gestructureerd interview en helpt de professional om kritisch te kijken naar het eigen oordeel. Binnen de RISC is er een vast onderdeel genaamd OXREC (Van den Berg et al., 2013). Dit instrument – ontwikkeld door de universiteit van Oxford – wordt gebruikt om algemene recidive en geweldsrecidive in te schatten. Middels de OXREC kunnen er inschattingen worden gemaakt op basis van de leefgebieden van een cliënt (zoals woonomstandigheden en delict historie) waarbij er voorspellingen uitkomen rondom het criminele gedrag van de cliënt. Uit deze OXREC komt een score, maar deze score wordt niet 1 op 1 binnen het model gepresenteerd aan de reclasseringswerker. Deze score wordt uiteindelijk gewogen door de professional die het afneemt. Het element OXREC is dus een AI-component Voor een uitgebreide validatiestudie van dit instrumentarium verwijzen wij graag naar de recente validatiestudie van Van den Berg en collega's (Van den Berg et al., 2021).

Ting en collega's (Ting et al., 2018) bekeken een AI risico-taxatie instrument voor de reclassering in Singapore om de kans op recidive te berekenen bij jeugddelinquenten. Dit risico-taxatie instrument zou een nauwkeuriger kans op recidive inschatten dan de andere aanwezige instrumenten binnen de reclassering van Singapore. Dit is echter een programma wat momenteel niet in gebruik is bij de reclassering in Singapore. De auteurs claimen dat de kennis vanuit dit programma bij zou kunnen dragen aan het creëren van interventies op maat voor de jeugddelinquenten, en zo hun kans op een betere toekomst te vergroten. De Singaporese reclassering heeft als taken om te beoordelen of gedetineerden onder reclasseringstoezicht kunnen komen te staan, en het voorzien in programma's voor degenen die onder toezicht staan (Ministry of social and family development Singapore, 2021). Er is een gedeelte van de organisatie die bestaat uit vrijwillige reclasseringswerkers. Voormalig gedetineerden staan tussen 6 maanden en 3 jaar onder toezicht. De Singaporese reclassering lijkt veel raakvlakken te hebben op het gebied van de structuur van de organisatie en de uit te voeren taken met de Nederlandse Reclassering (zie het rapport van de Ministry of social and family development Singapore, 2021). De resultaten uit het genoemde paper van Ting en collega's (2018) toont aan dat er net net als in Nederland interesse is in het toepassen van AI binnen de reclassering, maar nog niet op grote schaal wordt toegepast.

De papers die uit de scoping review komen tonen ons aan dat er weinig onderzoek wordt gedaan naar de toepassing van AI binnen het reclasseringswerk. Er wordt echter vooral gekeken naar hoe goed de AI zijn werk doet waarbij er grotendeels voorbij wordt gegaan aan de morele, ethische, juridische en praktische aspecten van deze systemen. Wat ook opvalt is dat in de papers waarin AI al toegepast wordt in de reclassering, de medewerkers die met deze systemen zullen moeten werken nog niet zijn betrokken in het onderzoek wat tot nu toe beschikbaar is.

### 4.1.3 AI binnen justitie en GGZ

Binnen het justitiële veld komen meerdere toepassingen naar voren op het gebied van AI. In tabel 6 (Appendix A) wordt een overzicht gegeven van de genoemde toepassingen. Hieronder behandelen wij de drie meest genoemde toepassingen: justitiële beslissingen, in het specifiek COMPAS (52 papers), predictive policing (16 papers) en AI als informatieclassificator (3 papers).

#### 4.1.3.1 Justitiële beslissingen en COMPAS

De meest genoemde toepassing van AI ligt in het nemen van beslissingen als risico-taxatie instrument. Alhoewel bijna alle 84 papers over AI binnen justitie dit onderwerp aansnijden, zijn er 52 papers die dit als hoofdonderwerp behandelen. AI wordt hierin toegepast om risico's te berekenen over bijvoorbeeld de kans op delinquentie.

Het meest gebruikte risico-taxatie instrument binnen het internationale justitiële veld is COMPAS (genoemd in 18 papers): Correctional Offender Management Profiling for Alternative Sanctions (Dressel & Farid, 2018). Dit programma wordt voornamelijk in de Verenigde Staten gebruikt om risico's op recidive in te schatten, waarna de rechter deze informatie gebruikt om gevangenisstraffen op te leggen. COMPAS kijkt naar een combinatie van factoren en berekent een score door middel van een algoritme (Deeks, 2018).

De voordelen van COMPAS – in de ondersteuning van justitiële besluitvorming in vergelijking met een rechter – liggen in de consistentie van altijd hetzelfde antwoord geven voor dezelfde omstandigheden alsmede ook de efficiëntie van de dataverwerking (Fortes, 2020). Een review van het systeem laat zien dat schaal een even accurate kan geven van het recidive niveau. Deze is vergelijkbaar met bekende risico-taxatie instrumenten (Brennan et al., 2009)

Er zijn uiteraard ook zwakheden binnen COMPAS. Verdachten wordt ontzegd hoe hun risicoscore is berekend (Deeks, 2018). COMPAS is van een commercieel bedrijf, welke onder andere door copyright claimde het algoritme niet te hoeven delen. Hierdoor kunnen verdachten niet de accuratesse en validiteit van het algoritme testen en bezwaar aantekenen. Er is ook niet na te gaan welke data is gebruikt om het algoritme te trainen. Dit is een kritiekpunt wat veelvuldig terugkomt als het gaat om het gebruik van dit programma, met name over de data gerelateerd aan etniciteit van de verdachten (Deeks, 2018; Dressel & Farid, 2018; Fortes, 2020; Moses, 2020). COMPAS gebruikt data die gecorreleerd zijn met etniciteit en mogelijk leiden tot discriminatie (Dressel & Farid, 2018; Fortes, 2020). Dit betekent dat etniciteit niet expliciet meegenomen wordt als factor, maar dat er factoren zijn (zoals bijvoorbeeld sociaaleconomische status, postcode) die in de weging van COMPAS een negatieve impact hebben waarbij die factoren vaker voorkomen bij zwarte Amerikanen.

In het geval van COMPAS lijken we zeker met een "black box" effect te maken hebben, waarin justitiële beslissingen door algoritmes gemaakt worden, maar wij niet kunnen weten waarop deze beslissingen exact zijn gebaseerd.

#### 4.1.3.2 Predictive policing

De tweede meest prominente toepassing van AI binnen justitie (16 papers) is de predictive policing. In Nederland en vele andere landen wordt gebruikt gemaakt van predictive policing (Hälterlein, 2021; Peeters & Schuilenburg, 2018). Bij predictive policing wordt gebruik gemaakt van algoritmes om te voorspellen waar criminaliteit waarschijnlijk plaats zal vinden. Op deze plekken wordt dan door de politie extra gesurveilleerd (Browning, 2021). Het gaat verder dan eerdere AI-programma's vanwege de voorspelling wat er in de toekomst zou kunnen gaan gebeuren (Peeters & Schuilenburg, 2018). Lee en collega's (2020) analyseerden algoritmes gebruikt bij predictive policing in Portland en Cincinnati in de Verenigde Staten en vonden een hoge efficiëntie en accuratesse.

Er zijn echter haken en ogen aan de inzet van de technologie, met daarbij vele tegenstanders. Alikhademi en collega's (2022) voerden recent literatuuronderzoek uit aan welke voorwaarden predictive policing moet voldoen voordat dit geïmplementeerd kan worden. Allereerst kwam naar voren dat er aandacht geschonken moet worden aan "biased data" en "fairness", oftewel bevooroordeelde data en eerlijkheid. De extra aandacht voor mogelijke criminaliteit op een plek in een stad kan ertoe leiden dat er een "self-fulfilling prophecy" ontstaat: juist door de extra aandacht wordt criminaliteit eerder opgemerkt (O'Donnell, 2022). Hierdoor zal de plek aangemerkt blijven als hoog in criminele activiteit, terwijl er voor de extra aandacht mogelijk evenveel criminaliteit was als een plek die niet aangemerkt werd door het algoritme. Dit kan vooroordelen in de hand werken (Browning, 2021). Het tweede punt van Alikhademi en collega's (2022) is dat er weinig aandacht is voor privacy and rechten en vrijheden van de burger. Door de toename van technologie in ons dagelijks leven, wordt de politie gedwongen om hier ook aan deel te nemen. Echter, het gebruik van

data door de politie om voor minder criminaliteit te zorgen, komt ook met de mogelijke impact op de privacy van burgers. Het derde punt van Alikhademi en collega's (2022) is het moeten tonen van transparantie en verantwoordelijkheid. Transparantie is een vereiste voor het behouden van vertrouwen tussen burgers en politie. Er moet openheid van zaken worden gegeven in hoe algoritmes werken en op welke data de uitkomsten van deze algoritmes gebaseerd worden. Dit gebeurt echter niet. Van belang is ook dat de politie verantwoordelijkheid wordt gehouden voor haar beslissingen, of deze nou beïnvloed zijn door een algoritme of niet. Bij wie de verantwoordelijkheid nou echt ligt voor slechte beslissingen door het algoritme – politie, justitie of de ontwikkelaars van de technologie – is niet duidelijk.

Voorstanders van predictive policing stellen dat deze technologie de efficiëntie van het justitiële systeem kan vergroten en – nog belangrijker – het justitiële systeem eerlijker kunnen maken (Shapiro, 2019). Dit omdat AI-systemen juist geen vooroordelen hebben zoals de mens deze heeft, maar puur opereert op basis van de data. AI-systemen zouden daardoor juist de vooroordelen tegen kunnen gaan en de inconsistenties in menselijke beslissingen verminderen. Een voorbeeld hiervoor komt van Shapiro (2019), waarbij agenten door een AI-systeem gevraagd werden om in bepaalde wijken te surveilleren. De politieagenten hadden op basis van hun ervaring niet gedachten om in deze wijken te surveilleren, maar deze surveillances bleken effectief en leidden tot nieuwe inzichten. Daarnaast droeg het ook bij aan het verminderen van verveling van de agenten, doordat ze van hun vaste surveillanceroutes afweken. De data waarop deze AI-systemen gebaseerd zijn zullen waarschijnlijk nooit geheel vrij zijn van vooroordelen. Brayne (2020) geeft aan dat professionele oordelen van politieagenten ook niet zonder vooroordelen zijn, en dat de combinatie van AI-systemen en professioneel oordeel van politieagenten – elk met enige mate van vooroordeel – effectief kunnen werken binnen predictive policing.

#### *4.1.3.3 Informatie classificatie*

De derde meest genoemde toepassing van AI binnen het justitiële veld (16 papers) is AI als informatie-classificator. Voor bibliothecarissen die justitiële informatie moeten verzamelen, analyseren en documenteren is er een overdaad aan informatie beschikbaar waarbij ook de taken die zij moeten uitvoeren aan het groeien zijn (Hilt, 2017). Door de inzet van AI-programma's kunnen bijvoorbeeld vragen over welke jurisprudentie relevant is sneller worden beantwoord, en wanneer deze informatie niet gevonden wordt kan er automatisch een bibliothecaris op de hoogte gesteld worden. Daarnaast kunnen de AI-systemen 24 uur per dag en 7 dagen per week antwoord geven op vragen (Hilt, 2017). Zulke AI-databases worden al ingezet binnen het justitiële veld door advocaten en rechters (Franssen & Berrendorf, 2021). Het snel vinden van de juiste informatie en het ordenen ervan is met het juiste AI-systeem een zeer waardevolle methode (Liu et al., 2019).

Uiteraard zijn ook hier nadelen aan het gebruik van deze systemen. Hilt (2017) benoemt hierin als eerste de mogelijke daling van informatie geletterdheid van de mens, de vaardigheid om de juiste justitiële stukken te vinden en te beoordelen. Als een AI-systeem deze vaardigheid grotendeels of geheel overneemt, blijft deze vaardigheid van de mens niet in stand. De vraag is ook in hoeverre het AI-systeem de finesse van een bepaalde zaak goed kan vangen, wat bij een bibliothecaris vaak wel het geval is. Het tweede nadeel is de kosten van deze programma's welke toenemen naarmate het systeem meer kan. Deze kosten kunnen zeer waarschijnlijk niet door alle potentiële gebruikers worden gedragen. Als laatste nadeel benoemt Hilt (2017) mogelijke problemen met de privacy. AI-systemen leren van eerdere zoekopdrachten. Het is mogelijk dat vertrouwelijke cliëntinformatie in handen van die systemen komt. Dit soort AI-systemen lijken niet het menselijk werk te gaan vervangen (Hilt, 2017) maar juist te ondersteunen.

#### *4.1.3.4 Overige toepassingen*

Naast bovengenoemde toepassingen zijn er ook drie papers die AI toepassen binnen de GGZ. Deze toepassingen hebben ook de functie van een risico-taxatie instrument. Daarnaast zijn er ook elf papers die een verscheidenheid aan AI-toepassingen laten zien. Dit gaat om bijvoorbeeld gezichtsherkenningssystemen om delinquenten sneller te herkennen, of het inzetten van chatbots bij delinquenten om vragen aan te stellen.

### **4.1.4 Aspecten in acht te nemen bij implementatie van AI**

Ongeacht hoe AI wordt ingezet, en in welke tak van het justitiële veld, zijn er duidelijke signalen over welke aspecten cruciaal zijn om in acht te nemen voor AI ingezet kan worden. Tabel 7 (Appendix A) geeft een overzicht van welke papers welke aspecten benoemen.



De meest belangrijke aspecten die voor toepassing van AI pleiten, is de mogelijkheid dat AI sneller en accurater beslissingen zouden kunnen nemen dan de mens (genoemd in 26 papers) en de mogelijkheid dat AI ook een onbevooroordeelde beslissing kan nemen (20 papers). Dit zijn uiteraard voordelen als een AI-systeem optimaal functioneert.

De meest belangrijke aspecten die genoemd worden als nadelen van AI-systemen – en die nauw met elkaar samenhangen – zijn de mate van bevooroordeeldheid van data en AI-systemen, het openen van de “black box” van de AI-systemen en de acceptatie van AI systemen door de gebruikers. Hieronder beschrijven wij deze nadelen van de implementatie van AI-systemen.

#### *4.1.4.1 Bevooroordeeldheid van data en AI-systemen*

Allereerst is er de eerder genoemde focus op “biased data” en “fairness”, oftewel bevooroordeelde data en eerlijkheid (Green, 2020). Dit punt is het meest cruciale punt om mee te nemen in het wel of niet implementeren van AI-systemen – wat ook te zien is in de 52 papers die dit onderwerp expliciet behandelen. Wat voorstanders van adviezen die gegeven worden of beslissingen genomen worden door AI-systemen stellen is dat AI-systemen vrij zijn van menselijke vooroordelen. De systemen opereren immers alleen op de data die in het systeem terecht komt, zonder aanzien des persoon (Beretta et al., 2019). Menselijke beslissingen en adviezen zijn nooit vrij van enig vooroordeel, hoe goed we dit ook tegen proberen te gaan en bewust van zijn. De toevoeging van AI-systemen in het nemen van beslissingen en/of het nemen van adviezen zou dus een welkome aanvulling of mogelijke vervanging van menselijke beslissingen zijn. Onderzoekers vinden ook dat er AI-systemen zijn die beter zijn in het nemen van beslissingen dan de mens (Beretta et al., 2019).

Echter, door het gebruik van bevooroordeelde data kan het zijn dat de eerlijkheid van de beslissingen van de AI-systemen niet altijd zo optimaal is als gewenst. AI-systemen leren van de data die wordt aangeleverd, en in deze data kunnen de vooroordelen verwerkt zitten die ook een rol spelen bij menselijke beslissingen (Beretta et al., 2019). Dit zorgt ervoor dat er goed gekeken moet worden naar de data die gebruikt worden door AI-systemen. Voor onderzoekers zijn er nu weinig middelen beschikbaar om de effectiviteit van de AI-systemen te kunnen testen en de eerlijkheid van deze uitkomsten en data die in de systemen terecht komen (Alikhademi et al., 2022). Oneerlijke data is moeilijk te identificeren (Eckhouse et al., 2019). Mogelijk ligt de oplossing in het verzamelen van meer persoonlijke data, welke echter wel sensitiever kan zijn (Žliobaitė & Custers, 2016). Een juridische basis voor welke data gebruikt kan worden en het reguleren van de eerlijkheid van deze AI-systemen ontbreekt (Demura & Klepka, 2021). Meer menselijk toezicht zal niet per definitie tot een reductie in de vooroordelen (Vaccardo & Waldo, 2019).

Er is nog werk te doen voordat we er zeker van kunnen zijn dat AI-systemen eerlijke beslissingen nemen door de onzekerheid van eerlijkheid van de data (Csányi et al., 2021; Vincent & Viljoen, 2020). Echter, er lijkt een nuance te missen in de discussie over de eerlijkheid van AI-systemen. Een directe vergelijking tussen de mate van vooroordelen van menselijke beslissingen en adviezen versus AI-systemen wordt nauwelijks gemaakt. Als AI-systemen vooroordelen hebben maar in mindere mate dan de mens, lijkt het eerlijker om deze AI-systemen in te zetten. Het lijkt onontkoombaar dat er enige mate van oneerlijkheid zal blijven in AI-systemen. We zullen moeten onderzoeken bij (toekomstige) AI-systemen en de adviezen en beslissingen welke hieruit komen in welke mate deze bevooroordeeld zijn of kunnen zijn, en deze continu blijven monitoren (Wright, 2020).

#### *4.1.4.2 Het openen van de black-box van AI-systemen*

Het gebrek aan toegang tot de algoritmes zelf en hoe deze precies werken is een kritiekpunt op alle AI-toepassingen binnen het justitiële veld als het aankomt op risicovoorspelling (Alikhademi et al., 2022). Dit wordt expliciet benoemd in 25 papers. De black-box is een feit bij deze toepassingen (Hälterlein, 2021), en we zullen hem moeten openen om beter te kunnen begrijpen hoe beslissingen genomen worden door algoritmes, en of de inzet van deze algoritmes verantwoordelijk is. Daarboven op komt het feit dat de AI-systemen ook van commerciële bedrijven kunnen zijn. Er zit een spanningsveld tussen deze commerciële bedrijven en de wensen van de gebruikers van deze systemen: de bedrijven willen winst maken en het algoritme niet prijs geven, terwijl de gebruikers graag toegang willen tot de data die gebruikt is voor deze beslissing (Carlson, 2017). Openheid van het algoritme binnen het AI-systeem als het aankomt op beslissingen die gemaakt worden – in het bijzonder van risico-taxatie AI – lijkt een cruciaal punt voor acceptatie van beslissingen en/of adviezen genoemd door AI-systemen.

#### *4.1.4.3 Acceptatie van AI-systemen door gebruikers*

Als laatste moet in acht worden genomen dat AI-systemen niet altijd geaccepteerd zullen worden door de gebruikers van het systeem: de personen die met de AI werken of afhankelijk zijn van de

uitkomsten (Scurich & Krauss, 2020; Swofford & Champod, 2021). Dit wordt expliciet benoemd in 13 papers. Mensen zijn vaak tegen het vervangen van de menselijke maat in het nemen van beslissingen gegenereerd door computers. Dit komt onder andere door het geloof dat mensen altijd beter zullen zijn in het maken van deze beslissingen. Dit is echter niet per definitie waar: AI kan beter presteren dan mensen (Bagaric et al., 2019; Vries et al., 2020). Moses (2020) benoemt dat we nu in een tijd zitten waarin de beslissingen die genomen worden door AI-systemen binnen het justitiële veld, een belangrijke impact hebben op hoe mensen ze waarnemen en hoe deze verder ontwikkeld worden. Hoe dit te onderzoeken is een volgende stap.

#### 4.1.5 Conclusie scoping review

Vanuit deze scoping review zijn er signalen dat AI potentie kan hebben om ingezet te kunnen worden binnen 3RO. In antwoord op onderzoeksvraag 1 (de meest toepasselijke definitie van AI) zien we dat er geen algemeen geldende definitie is. De kern ligt echter wel in (het ontwerpen van) intelligente systemen die net zo veel of nog meer kunnen dan mensen.

In antwoord op onderzoeksvragen 3 en 4 (de huidige toepassing van AI binnen het reclasseringswerk en kansen voor toepassing van AI) zien we dat er internationaal nog weinig gebruik wordt gemaakt van AI binnen het reclasseringswerk. Binnen het justitiële veld en de GGZ (onderzoeksvraag 1) zien we het meest prominente gebruik van AI binnen het justitiële veld als risico-taxatie instrument. Risico-taxatie instrumenten bestaan in het justitiële veld al zeer lange tijd – alsook binnen 3RO – en zullen ook gebruikt blijven worden in de toekomst (Harkens, 2018; Werth, 2019). Met de komst van AI-systemen kunnen deze beslissingen en adviezen op basis van risico-taxaties mogelijk sneller worden genomen (Sil & Roy, 2019). Een andere belangrijke toepassing die voor 3RO waardevol kan zijn – waar waarschijnlijk minder aandacht voor is – is het gebruik van AI als informatie-classificatie systeem. AI-systemen kunnen helpen met het ordenen van data en de afhandeling hiervan (Ruhl et al., 2018). Gegeven de grote hoeveelheid informatie waarmee reclasseringswerkers te maken hebben in hun dagelijks werk, zal dit kunnen bijdragen aan hun efficiëntie.

In antwoord op onderzoeksvragen 5 en 6 (voor- en nadelen van AI binnen het reclasseringswerk en welke aspecten in acht te nemen) vinden we dat wanneer AI-systemen goed functioneren deze potentieel werk uit handen nemen van reclasseringswerkers. Echter, duidelijk is dat er nog veel haken en ogen zitten aan de inzet van AI-systemen door de mogelijke bevooroordeeldheid, het black-box effect en de acceptatie van AI-systemen door gebruikers.

AI kan zeker in het juridische veld nooit geheel de mens vervangen (Boahemaa, 2019; McKay, 2020), alhoewel het niet waarschijnlijk is dat dit ooit zal gebeuren (Morison & Harkens, 2019). Er is veel maatwerk nodig wat binnen AI-systemen problematisch is en waarschijnlijk ook zal blijven (Giannini, 2021). De vier producten die 3RO levert (advies, toezicht, werkstraffen en justitiële interventies) vereisen zo accuraat mogelijke oordelen vanwege de impact die de beslissingen op basis van deze oordelen kunnen hebben.

AI in het juridische veld zal echter niet zomaar verdwijnen, maar het lijkt juist meer aandacht te krijgen en zich sterk te ontwikkelen in de komende jaren (Docarmo et al., 2021). De beslissingen die door de AI-systemen genomen gaan en kunnen worden, moeten echter nog wel ontwikkelingen doormaken voordat zij volledig ingezet kunnen worden (Duker & Lodder, 1999). We zullen kritisch moeten zijn en blijven op de inzet van AI-systemen (Fortes, 2020; Peeters & Schuilenburg, 2018) en de systemen continu blijven ontwikkelen als deze ingezet worden (Wright, 2020).

## 4.2 Resultaten interviews met experts

Op basis van de gehouden interviews behandelen we de resultaten van de interviewstudie in vier thema's:

1. Definitie van AI (onderzoeksvraag 1)
2. Toepassingen van AI binnen reclassering en het justitiële veld (onderzoeksvraag 1,3, en 4)
3. Morele, ethische en juridische aspecten aangaande toepassing van AI binnen 3RO (onderzoeksvraag 5 en 6)
4. Stappen nodig om AI binnen 3RO succesvol in te zetten (onderzoeksvraag 7 en 8)

Hieronder beschrijven wij per thema hoe de experts hierover dachten. Uit de interviews halen wij quotes ter illustratie.

### 4.2.1 Definities van AI

Onze bevindingen uit de scoping review worden bevestigd: er is geen meest gebruikte definitie van AI. Er komen verschillende opties ter sprake om AI te definiëren.

---

*"Dus ik zat een tijdje terug in een werkgroep van de stichting toekomstbeeld der techniek en daar kwamen we er eigenlijk ook niet uit wat AI nou precies is. En toen had ik geopperd: ja, AI dat is wat AI onderzoekers bezighoudt." (dr. van Maanen)*

*"Kunstmatige intelligentie dus. Dat betekent voor mij dat een computersysteem in hele grote hoeveelheden data zoekt naar correlaties en op basis daarvan mensen kan helpen om de wereld op een bepaalde manier te interpreteren." (prof. dr. Verbeek)*

*"Als je het aan tien verschillende mensen vraagt, krijg je tien net iets andere definities. Maar één hele algemene die volgens mij het meest wordt gebruikt, is een oude definitie van John McCarty uit 1956, ik weet niet of jullie die al hebben zien langskomen op jullie onderzoek. Het is zoiets als the science and engineering of making intelligent machines." (prof. dr. Bosse)*

---

Er werd door prof. dr. Bosse en dr. van Maanen een vergelijking gemaakt met de vroegere AI. In de vroegere AI werd de intelligentie in een systeem geprogrammeerd met behulp van codeerde regels geïmplementeerd door een programmeur, waarbij op basis van deze regels een voorspelling werd gedaan omtrent een bepaald onderwerp.

---

*"Dat wordt ook wel Weak AI of Narrow AI genoemd, dus het kan heel goed gezichten herkennen of plaatjes van honden en katten herkennen of een spelletje schaak spelen. Eén specifieke taak kan het heel goed, maar het kan niet zo even iets heel anders gaan doen. Een algoritme dat geleerd heeft om de kat te herkennen, kan niet vervolgens ook een potje schaken. Daar moet je weer een heel nieuw systeem voor bouwen zeg maar. En dat hebben wij mensen wel, we hebben een soort algemene kennis en we begrijpen ook wat we doen." (prof. dr. Bosse).*

---

Bij de moderne AI vormt en traint het systeem de regels zelf middels de data die het krijgt als input, en deze systemen zijn dus veel intelligenter in vergelijking met de vroegere AI. Dit betekent dat de definitie van wat precies AI is verschuift in de loop van der tijd:

---

*"Het is een soort verschuivende horizon." (dr. van Maanen)*

---

Hieruit kunnen wij opmaken dat er zelfs onder experts nog geen consensus is over wat AI is, maar het is duidelijk dat het gaat om intelligente systemen die ontzettend veel kunnen en dat het in ieder toepassingsdomein mogelijkheden kan bieden.

Prof. dr. Bosse geeft verder aan dat het belangrijk is om onderscheid te maken tussen Explainable AI, Responsible AI en Socially Aware AI (voor een verdere uitleg over deze termen verwijzen wij naar pagina 11 waarin deze termen verder uitgelegd zijn).

### 4.2.2 Toepassingen AI binnen reclassering en het justitiële veld

Er komen bij de experts geen voorbeelden van toepassing van AI binnen de reclassering zowel nationaal als internationaal ter sprake. Dit ondersteunt de resultaten van de scoping review waarin er maar een beperkt aantal toepassingen van AI binnen de reclassering werden gevonden op (inter)nationaal niveau.

Voor voorbeelden binnen het justitiële veld wordt verder verwezen naar predictive policing en AI systemen die rechters ondersteunen bij het nemen van beslissingen. Aangestipt wordt dat AI binnen het justitiële veld al een groot domein is en het nog meer toepassingen zal hebben in de (nabije) toekomst.

### 4.2.3 Morele, ethische en juridische aspecten aangaande toepassing AI binnen 3RO

De experts waren voornamelijk geïnteresseerd in de morele, ethische en juridische aspecten aangaande van de toepassing van AI binnen 3RO en het justitiële veld in zijn geheel.

Allereerst werd er gesproken door eenieder over de biases en fairness van AI die ook al aangestipt werden in de scoping review.

---

*"Maar om het acceptabel te maken, moeten we denk ik gewoon heel goed weten dat er met name die bias er niet in zit. Dat de data dus betrouwbaar is en representatief is." (prof. dr. Bosse)*

*"Het kernwoord is natuurlijk heel vaak fairness, dus dat moet wel rechtvaardig zijn." (prof. dr. Verbeek)*

*"Omdat dus gewoon blijkt dat als je te veel vertrouwt op data dat je dan dus ook gebiasde output krijgt. En dat is een groot probleem." (dr. van Maanen)*

---

Ongeacht wat voor soort AI-systemen worden ontwikkeld en toegepast, blijft dit een kritiekpunt. Er wordt vaak vergeten dat de uitkomst erg afhankelijk is van de precieze data waarop getraind is. Als er biases in de data zitten die wordt aangeleverd aan het AI-systeem, dan worden deze biases vaak nog een keer uitvergroot door een algoritme in de uiteindelijke uitkomst. Dit kan bijvoorbeeld het geval zijn bij predictive policing: als er toevallig in een dataset data zit die weergeeft dat er op een bepaald moment de meeste misdaad is voorgekomen in multiculturele wijken, dan zal een algoritme daarvan leren en dus voorspellen dat in die wijk extra gesurveilleerd moet worden terwijl dat mogelijk niet legitiem is.

---

*"Er zitten dus veel bronnen van onzekerheid in die de uiteindelijke uitkomst kunnen beïnvloeden." (dr. van Maanen)*

---

Alleen als biases geëlimineerd en/of gereduceerd wordt en fairness wordt bevorderd, kan het AI-systeem optimaal functioneren. Zeker binnen 3RO is het bereiken van deze balans cruciaal, gegeven de (soms beladen) beslissingen die genomen moeten worden door de reclasseringswerkers.

---

*"Dus leerprocessen, denk ik, lijkt mij in eerste instantie het allerbelangrijkste in deze context." (prof. dr. Verbeek).*

---

In aansluiting hierop moet ook gekeken worden naar de culturele context van de beslissingen die genomen worden en hoe AI-systemen hiermee omgaan. Als AI-programma's toegepast worden die in een andere culturele context zijn ontwikkeld dan in Nederland, kan dit gevolgen hebben voor de besluitvorming.

---

*"Ja, ik denk in ieder geval dat dat aandachtspunten zijn van: kunnen er vooroordelen inzitten als je data gebruikt uit het verleden, waardoor sommige mensen er meer slachtoffer van zijn dan anderen, en kun je zorgen dat die uitleg goed is?" (prof. dr. Bosse)*

*"En dat maakt het dan meteen wat ingewikkelder, want dan is eigenlijk de vraag hoe AI-systemen op zo'n manier een duiding van de wereld maken of een advies kunnen geven wat je ook ruimte biedt aan die culturele variëteit." (prof. dr. Verbeek)*

---

Tot slot werd er door de experts aangekaart dat er bewust nagedacht moet worden over de privacy van de reclassanten en de verkregen data voortkomend vanuit de casussen, mede als het waarborgen van de veiligheid van deze data.

---

*"En iets anders is nog het hele verhaal rondom privacy en security, dus die data die komt ergens vandaan, die moet ergens opgeslagen worden. Nou ja, wie kan daarbij?" (prof. dr. Bosse)*

*"Dat zou voor mij een aandachtspunt zijn, dat dat systeem zo gedifferentieerd is dat je niet meer per definitie – omdat het nou eenmaal het systeem is dat moet werken – al die gegevens moet verzamelen. Maar alleen de gegevens die relevant zijn om deze personen te kunnen beoordelen." (prof. dr. Verbeek)*

---

---

---

Een eenduidig antwoord of AI binnen 3RO verder ingezet moet worden, is dus niet makkelijk te beantwoorden.

---

*"Want ja, het gaat om een vonnis, iemands leven, vaak iemands vrijheid ontnemen en ja, wat voor beslissingen kun je dan overlaten aan een computer? Of wat moet echt door een mens gebeuren? Nou, dat zijn grote keuzes, denk ja." (prof. dr. Verbeek).*

*"Veel algoritmes zoeken toch naar een soort generalisatie, terwijl individuen toch echt wel verschillen. En het is natuurlijk super frustrerend als jij wordt afgerekend omdat je heel erg lijkt op andere personen die eigenschap x, y, z hebben." (dr. van Maanen)*

---

#### 4.2.4 Stappen nodig om AI binnen 3RO succesvol in te kunnen zetten

Er wordt door de experts erkend dat het toepassen van AI-systemen met voordelen kan komen voor de reclasseringspraktijk. Er kan eventueel meer maatwerk worden geleverd doordat AI-programma's meer inzicht kunnen geven in patronen uit het verleden en heden, en welke voorwaarden er van toepassing zijn om verantwoord terug te kunnen keren in de samenleving. Dit kan effect hebben op bijvoorbeeld (het inschatten van) recidivisme of op de samenwerking tussen cliënt en reclasseringswerker. Daarnaast kan er ook efficiënter gewerkt worden middels het gebruik van AI-systemen door een reclasseringswerker.

---

*"Zo'n systeem kan in theorie duizenden casussen tegelijk doornemen terwijl je als mens dat allemaal met de hand moet doen." (prof. dr. Bosse).*

---

Daaropvolgend zijn er door de experts verschillende stappen benoemd die nodig zijn om AI binnen 3RO – en daarmee ook het justitiële veld als geheel – succesvol in te kunnen zetten.

Allereerst moet er een periodieke check van het systeem komen om te kijken of de fairness van het systeem nog optimaal werkt. De algoritmes en bijbehorende data moeten dus up-to-date blijven en niet ongewijzigd worden toegepast. Bias is onvermijdelijk, maar we zullen elke keer moeten blijven leren van de AI-programma's om ons heen om deze bias zoveel mogelijk te reduceren. AI moet dus allereerst Responsible AI zijn. Ongeacht waar en hoe AI wordt ingezet, is dit de eerste prioriteit.

---

*"Dus je ontkomt er niet aan om mensen te profileren, daar gaat het juist om. Maar dat op een verantwoorde manier doen is ongelooflijk belangrijk. Dus in dit geval is natuurlijk denk ik het kernpunt hoe kun je ervoor zorgen dat dat systeem een betrouwbare voorspelling geeft." (prof. dr. Verbeek)*

---

Prof. dr. Verbeek benoemt Hybrid Intelligence: de hybride vorm van besluitvorming waarin mensen samen met computersystemen keuzes maken. Wanneer kies je voor de besluitvorming door het algoritme, en wanneer de besluitvorming van de mens? Onze kennis van deze hybride besluitvorming is nog in een beginstadium, maar het duidt op een ondersteuning van de reclasseringswerker bij het nemen van beslissingen middels het AI-systeem. Hierin speelt ook de Explainable AI weer een rol. Als AI-systemen een goede uitleg kunnen geven over waarom een besluit genomen wordt, kan er beter gereflecteerd worden of het besluit van het algoritme of het besluit van de mens gevolgd moet worden.

---

*"Het systeem moet op een moreel verantwoorde manier in elkaar zitten, maar de inbedding van het systeem in de praktijk waarin mensen een rol spelen moet ook moreel verantwoord in elkaar zitten." (prof. dr. Verbeek)*

*"In veel gevallen zou ik dat ook adviseren. Ja, er zijn misschien mensen die zeggen dat het niet meer nodig is, maar in de meeste gevallen is dat toch wel een soort extra veiligheid." (prof. dr. Bosse)*

*"Ik denk dus dat een generaliserend algoritme op zich ook fair kan zijn, maar dan zou ik denk ik niet zo snel een beslissing grotendeels aan een algoritme overlaten. Dan zou ik het meer als een decision-support zien met belangrijke functies zoals een explainable component erin. Hoe persoonlijker hoe betrouwbaarder zou je ook kunnen zeggen." (dr. van Maanen)*

---

Belangrijk is ook dat de gebruikers van de AI-systemen – zowel cliënten als reclasseringswerkers – gehoord moeten worden over de toepassing van deze systemen. Prof. dr. Bosse vertelt over zijn studies waarin hij heeft gekeken naar de acceptatie van AI-systemen onder verschillende omstandigheden.

---

*"Er blijkt eigenlijk iedere keer uit dat het voor de meeste mensen eigenlijk niet zoveel uitmaakt als dus de kwaliteit van het advies goed genoeg is, dan hebben ze geen voorkeur voor de mens of een algoritme bij het uitbrengen van een advies."* (prof. dr. Bosse)

---

Wanneer de reclasseringswerkers de AI-systemen niet accepteren, kan de implementatie hiervan nooit succesvol zijn.

Daarnaast moet er een begrip zijn van wat AI doet en waar het mis zou kunnen gaan.

---

*"Op die manier creëer je een soort gevoel van vertrouwen in dat systeem die, nou ja misschien niet het beste met je voort heeft want dat is dan weer een hele menselijke eigenschap, maar in ieder dat het systeem jou kan helpen om bepaalde doelen te bereiken."* (dr. van Maanen)

---

Een gesuggereerde toepassing van AI binnen 3RO is het gebruik van chatbots:

---

*"Dus dat gaat over bijvoorbeeld robots, sociale robots die in ziekenhuizen steeds meer vindt of op scholen of chatbots, dat zijn van die computerprogramma's waar je via tekst mee kunt spreken. Dus als je op een website bent van de IKEA of zo die je dan vraagt, wil je dit kopen en dan kun je daar mee in gesprek gaan. Dus dat soort technologie. Ik kan me op zich voorstellen dat er in het domein van de reclassering mogelijk ook iets met dat soort technologie gedaan zou kunnen worden"* (prof. dr. Bosse)

---

Er zijn dus vele stappen die te nemen zijn in het succesvol implementeren van verdere AI-systemen binnen 3RO.

#### 4.2.5 Conclusie expert interviews

Wederom wordt duidelijk dat AI een groot onderzoeksveld is en dat er vele definities zijn over wat AI nou precies inhoudt (onderzoeksvraag 1). De experts hadden minder weet over toepassingen van AI binnen het justitiële veld en het reclasseringswerk, maar konden duidelijk aangeven waar de mogelijke valkuilen liggen in het toepassen van AI binnen 3RO (onderzoeksvraag 1, 3 en 4).

Zoals ook uit de scoping review blijkt, zitten er vele haken en ogen aan de inzet van AI-systemen. De meest belangrijke blijft de mogelijke bias in data en de toepassing van algoritmes, en de daaraan gerelateerde fairness van het AI-systeem. Alhoewel hier nog geen kant-en-klare oplossing voor is, is het duidelijk dat AI-systemen constant bekeken moeten worden in welke mate ze bias vertonen (onderzoeksvraag 5 en 6). Dit is voor 3RO in het bijzonder een cruciaal, gegeven de impact van de beslissingen op het dagelijks leven van de cliënten.

Het betrekken van reclasseringswerkers in vervolgonderzoek(en) zal nodig zijn om medewerkers te overtuigen om met AI-systemen (samen) te werken en de relevantie hiervan in te zien. Hybride Intelligence – de hybride vorm van besluitvorming waarin mensen samen met computersystemen keuzes maken – zal dus ook nader onderzocht moeten worden. Wanneer kan een reclasseringswerker er van uit gaan van zijn of haar beslissing, en wanneer is het nodig om de besluitvorming (gedeeltelijk of in zijn geheel) te leggen bij een algoritme?

De experts doen weinig uitspraken over op welke wijze AI binnen 3RO verder toegepast kan worden buiten de al bestaande RISC. Geopperd wordt om na te denken over chatbots die (basale) vragen van cliënten kunnen beantwoorden. Dit veld is echter nog sterk in ontwikkeling en kan dus nog niet complexe vragen beantwoorden die binnen 3RO wel nodig zijn (onderzoeksvraag 7 en 8).

Het toepassen van AI binnen 3RO is uiteraard mogelijk, maar er zijn dus vele randvoorwaarden waar over nagedacht moet worden voordat deze implementatie succesvol is.

### 4.3 Resultaten interviews met reclasseringswerkers

Op basis van de gehouden interviews met reclasseringswerkers en de onderzoeksvragen komen er vier thema's naar voren uit het interview:

1. De kennis van AI (onderzoeksvraag 1 en 2)
2. Het huidige gebruik van AI binnen 3RO (onderzoeksvraag 2 en 3)
3. Mogelijke toepassingen van AI binnen 3RO (onderzoeksvraag 4, 5 en 6)
4. Stappen nodig om AI succesvol in te kunnen zetten binnen 3RO (onderzoeksvraag 7 en 8)

Hieronder beschrijven wij voor elk thema hoe de reclasseringswerkers hierover dachten. Uit verschillende interviews halen wij quotes ter illustratie waarbij de bijdrages genummerd zijn per geïnterviewde medewerker.

#### 4.3.1 Kennis van AI

Duidelijk wordt dat de meeste reclasseringswerkers al gehoord hebben van AI, voornamelijk in de context van hun dagelijks leven. Deze ervaringen zijn niet altijd positief.

---

*"Dus niet in de uitwerking, maar wel dat het op allerlei manieren ingezet en gebruikt wordt en tot ergernissen leidt, omdat ik dan ergens iets gezocht heb en vervolgens doodgegooid wordt met allerlei onzin. Dan denk ik, ik heb al lang een wasrek, bijvoorbeeld, hé, ja, dan wil ik niet nog weer een wasrek zien" (reclasseringswerker 3).*

*"Ik vind het AI in m'n privéleven ook wel een beetje een naar of soms een vervelend bijmaakje hebben. Ik heb bijvoorbeeld laatst een nieuwe auto gekocht. Nou, ik kan nu m'n Facebook niet openen of ik zie de auto reclames nog steeds voorbijkomen als ik een YouTube filmpje open, dan krijg ik reclame over auto's en zo vreselijk." (reclasseringswerker 4).*

---

De presentatie met informatie over AI en de toepassingen van AI in 3RO – die van te voren werd gestuurd naar de reclasseringswerkers en die door eenieder was bekeken van te voren – werkte daarom ook verhelderend. Zoals reclasseringswerker 2 aangaf:

---

*"Voordat ik die presentatie zag, dan denk ik eerder iets aan robots en dat soort techniek. Dan dacht ik niet als eerste aan een Siri of dat soort dingen. En door die presentatie dacht ik van oh, ja, dat zijn natuurlijk de linken met wat er in je dagelijks leven daar al mee gebeurt" (reclasseringswerker 2)*

---

Toch zijn de mogelijkheden die AI biedt niet voor een ieder duidelijk.

---

*"Een algoritme is principe gewoon een serie opdrachten na elkaar. Zit er in AI iets wat zelflerend is? Of is het juist AI als er iets zelflerend in zit? Of is enkel een programma met algoritmes al AI." (reclasseringswerker 1)*

---

Dit wijst erop dat alhoewel AI voor de meesten een redelijk bekend fenomeen is, er niet een alomvattend begrip is wat AI kan doen.

#### 4.3.2 Huidig gebruik van AI binnen 3RO

In de presentatie die van te voren werd bekeken door de geïnterviewde reclasseringswerkers werden de systemen die momenteel door 3RO gebruikt benoemd. Allereerst is er het eerder beschreven risicotaxatie-instrument, de RISC. Binnen de RISC is er een vast onderdeel genaamd OXREC (Van den Berg et al., 2013). De OXREC opereert op basis van AI en wordt gebruikt om algemene recidive en geweldsrecidive in te schatten. Er worden inschattingen gemaakt op basis van de leefgebieden van een cliënt (zoals woonomstandigheden en delict historie) waarbij er voorspellingen uitkomen rondom het criminele gedrag van de cliënt. Uit deze OXREC komt een score, welke afgewogen wordt door de reclasseringswerkers samen met de informatie die verder ingevuld is in de RISC. Aanvullend op de algemene risicotaxatie zijn er modules die gebruikt kunnen worden – wederom ook AI-aanvullingen – die voor specifieke delicten gebruikt worden, zoals bij stalking of zedendelicten. Zoals de SSA voor zedendelinquenten uitgelicht door reclasseringswerker 3:

---

*"De SSA is dus een instrument dat bestaat uit drie onderdelen. De S staat voor Static, de tweede S staat voor Stable en de derde staat voor Acute. En de Static is eigenlijk historische vragen. Volgens mij zijn het er tien die gaan over onveranderbare zaken alziende je aantal veroordelingen nou, je leeftijd, geslacht [...]. De Stable is eigenlijk ook wel wat lastiger te veranderen, maar kan wel. Maar daar zitten ook vragen in die meer tijd en energie zouden vergen om te kunnen veranderen. Bijvoorbeeld zit je sociaal netwerk uit een aantal goeie mensen of niet [...]. En de Acute is eigenlijk op dit moment even een vragenlijst of er risico's zijn of niet, die in principe elke maand afneemt en die is vrij snel met een vraag of vijf te beantwoorden. En de totale score van die instrumenten die geven indicatie tot welke risicogroep je dan behoort."*

---

Ieder van de zes (voormalig) reclasseringswerkers heeft met de RISC gewerkt, maar niet iedereen was zich er bewust van dat de RISC ook onder AI valt.

---

*"Ja, nou, ja, en de RISC bijvoorbeeld. Ook vond ik ook wel dat ja, dat weet je ergens wel, maar je koppelt niet per definitie daar een ja, een geprogrammeerd ding achter, weet je wel. Dat is gewoon oh, ja, de RISC die doe je gewoon punt en er wordt niet zoveel over nagedacht, zeg maar" (reclasseringswerker 5).*

---

Dit bewustzijn kwam enkel pas na het zien van de vooraf gestuurde presentatie.

Over het gebruik van de RISC zijn de reclasseringswerkers redelijk tevreden.

---

*"Goed. Noodzakelijk. Dat zorgt er wel voor dat dat je goed moet nadenken van wat je zelf vindt eigenlijk. En zeker ook als het programma afwijkt van wat je zelf denkt dan is dat interessant. Dan gaat er ergens iets niet goed. Dat kan aan alle kanten zijn. Maar het is goed en dat is ook om ons werk te kunnen verantwoorden." (reclasseringswerker 1)*

*"Nee, ik weet dat we wel eens discussies hebben, inderdaad van wat scores die d'r uit komen en het afwijken en wat dat doet met de betrouwbaarheid. Omdat je professioneel oordeel ook gekleurd is door wat je ziet, maar ook wel een welkome aanvulling is. Daar zit een spanningsveld in." (reclasseringswerker 2)*

---

De RISC wordt ervaren als een goed ondersteunend AI-programma waarbij het belangrijk is dat de reclasseringswerker het eindoordeel heeft.

---

*"Nou ja, je wordt een beetje gestuurd, maar wel op basis van onderzoek en weet je het is wel wetenschappelijk onderbouwd dat er achter zit, weet je wel. Dus ik vond het zelf altijd een beetje arrogant van ons, van mezelf ook hoor, maar van oh, ja, maar zo'n RISC wat zegt dat nou. Weet je, ik kan toch veel beter inschatten als ik met iemand in gesprek ben, hé hoe iemand erbij zit en wat voor risico's zou zijn. Terwijl, ja, je bent ook maar een mens, je hebt ook blinde vlekken en zo'n systeem, ja, dat is ook bewezen dat dat dus gewoon bijdraagt aan een betere inschatting dus ja, ik vind het super dat dat er is. Het is juist echt een meerwaarde." (reclasseringswerker 5)*

*"Nee, het is niet alleen instrument. He. Nee, het is niet zo dat je alleen dat instrument gebruikt en dan denkt: klaar. Maar dat is één. Maar goed ook, daar geldt natuurlijk dat als je denkt van hé, ik twijfel. Ik wil iemand mee laten kijken. Dat kan natuurlijk altijd." (reclasseringswerker 3)*

*"Nee, ik vind op zich wel een helder instrument. We hebben d'r ook een uitgebreide training in gehad om voordat we dat instrument mochten gebruiken. Er zit een duidelijke handleiding bij." (reclasseringswerker 4)*

---

Er zitten echter ook nadelen aan. Eén ervan is de tijd die het inneemt om de RISC in te vullen.

---

*"Alleen het is gewoon verschrikkelijk om in te vullen, hoor ik om me heen de hele tijd. Iedereen haat de RISC dus ja, ja, dat is wat het is." (reclasseringswerker 5)*

*"Het grote probleem is dat we dus standaard een RISC moeten invullen en dat is in de ogen van reclasseringswerker een hoop gedoe, een hoop werk en dat moet allemaal heel snel." (reclasseringswerker 6)*

---



Een ander minpunt is dat niet alles te vangen is met een AI-systeem.

---

*"Ja, nou, het is zeker ook dat ik dan zeg: ja, met zoals die zedendelinquent... dat klopt dat weer niet helemaal. Iemand die een lange crimineel bestaan heeft geleid, waar een hoog risico uit komt, terwijl die misschien nu gewoon huisje, boompje beestje heeft, werk en noem maar op. Het risico blijft hoger en daar duidt zo'n OXREC dan weer op van: ja, ziet er nu allemaal wel goed uit, maar iemand zijn flex is jarenlang geweest om crimineel bestaan te hebben met drugs en diefstal. Er hoeft maar genoeg met mis te gaan om daar weer in terug te vallen."* (reclasseringswerker 1)

---

Ook is er geen mogelijkheid om de RISC aan te passen met ervaringen vanuit de praktijk.

---

*"Ja, ik kan die informatie dan nergens kwijt, hé. Die informatie van openstaande zaken of zo, die kan ik niet ergens kwijt. Dus dan weegt het instrument het ook in geen enkel opzicht mee. Dus dan kan ik het alleen maar kwijt in m'n professionele oordeel en dan beschrijf ik dat gewoon."* (reclasseringswerker 4).

---

Interessant is dat er uit de interviews blijkt dat op het internationale vlak wel al gekeken wordt naar de toepassing van AI binnen de reclassering. Deze informatie is zeer waarschijnlijk niet in de scoping review naar voren gekomen omdat het niet-gepubliceerde resultaten betreft van een congres. Het betreft een pilot waarin door middel van gestandaardiseerde teksten cliënten gestimuleerd werden om zijn of haar doelen te behalen.

---

*"Dit congres stond ook onder andere in het teken van Artificial Intelligence dus daarom sloeg ik hier ook op aan dat ik dacht oh. Ja, daar heb ik ook wel een aantal voorbeelden gehoord vanuit Amerika, hoe ze nou eigenlijk [...] nou programma's hebben ontwikkeld om cliënten bijvoorbeeld te motiveren door bepaalde gestandaardiseerde teksten op te sturen, maar die dan bijvoorbeeld wel heel erg specifiek op de doelen van de cliënt zijn gericht. [...] Ze hadden het ook over dat het slagingspercentage daardoor van de behandeling ook veel meer toenam. ik weet niet in hoeverre ze dat hebben getoetst aan een andere, zeg maar zo'n controlegroep."* (reclasseringswerker 5)

---

Dit wijst erop dat er mogelijk meer vorderingen gemaakt worden in het toepassen of pogen toe te passen van AI binnen de internationale reclassering. Gekeken naar de presentaties die op het congres werden gehouden (Technology in Probation Expert Meeting in Helsinki, welke plaatsvond in 2018), was de focus van dit congres niet AI in de reclassering maar de toepassing van technologie in het algemeen. Niet elke presentatie is vrij beschikbaar of toont genoeg informatie die bruikbaar is voor dit rapport. Duidelijk wordt dat de Verenigde Staten (zoals ook blijkt uit de scoping review) de leiding nemen in het gebruik van technologie en AI binnen de reclassering, en dat er weerstand lijkt te zijn om AI en andere soorten technologie in het reclasseringswerk in te zetten. Er is dus binnen de Europese reclasseringsorganisatie wel interesse in en behoefte om te onderzoeken wat technologie in hun organisatie kan bijdragen.

#### 4.3.3 Mogelijke toepassingen AI binnen 3RO

Vanuit de scoping review en interviews met experts zijn er 3 potentiële mogelijkheden naar voren gekomen om AI binnen 3RO toe te passen. Deze mogelijkheden zijn:

- Verdere uitbreiding en inzet van AI-gebaseerde risicotaxatie-instrumenten, zoals de RISC.
- AI als informatieclassificator, in bijvoorbeeld de ordening van relevante informatie binnen documenten.
- AI in communicatie met cliënt, zoals bijvoorbeeld chatbots.

Elk van deze mogelijkheden hebben wij besproken en uitgevraagd over het nut en de noodzaak bij de reclasseringswerkers.

##### 4.3.3.1 AI in risico-taxatie instrumenten

Alle reclasseringswerkers zijn het erover eens dat ze geen verdere uitbreiding willen van beslissingen door AI binnen risicotaxatie-instrumenten. De reclasseringswerkers zijn het er unaniem over eens dat bij de RISC of toekomstige risicotaxatie-programma's het eindoordeel moet blijven liggen bij de reclasseringswerker.

---

*"Ik denk dat dat niet positief ontvangen gaat worden als er wel hoogopgeleide eigenwijze professionals die daar ook gewoon zelf graag een eindoordeel over willen vellen. En ik denk dat dat ook goed is. Die AI echt wel ondersteunend kan zijn, maar ik aarzel wat bij het hebben van een eindoordeel over iets wat dan ook direct zou leiden tot een bepaald advies." (reclasseringswerker 2)*

*"Ik denk bij ons in het werk dat het belangrijk is... Je hebt zoveel variabelen die ook niet allemaal meetbaar en concreet te maken zijn. Bijvoorbeeld ook wat betreft een inschatting maken of iemand zich onttrekt aan voorwaarden ja of nee. [...]Het gaat er ook om wat voor klik je met iemand hebt. En is dat willekeur? Ja, het is willekeur of je de klik hebt tussen twee personen als een toezichthouder en een cliënt. En dat laten inschatten door een programma lijkt mij niet echt mogelijk." (reclasseringswerker 1)*

*"Ja, ik heb liever dat dat minder is. Het blijft een computer, het blijft een niet mens. Dus ja, ik vind het niet helpend als dat systeem nog veel meer gaat schrijven." (reclasseringswerker 4)*

---

Wel zou het gewaardeerd worden als er in het gebruik van de RISC inzicht komt in hoe het AI-programma tot een bepaalde keuze komt. *"Ja, en ook de uitkomsten meer onderbouwd op verschillende leefgebieden zeg maar. Dat er niet alleen gezegd wordt risico laag, midden, hoog, maar ook hoe het programma daartoe komt precies." (reclasseringswerker 1)*. Momenteel kan het programma dit nog niet uitleggen. Dit wijst erop dat Explainable AI een goede aanpassing zou kunnen zijn voor de RISC.

Bij een verdere inzet van AI als risicotaxatie-instrument moeten er dan wel vaker checks gedaan worden of de informatie die in het programma is gestopt nog klopt.

---

*"Er zit wel een risico in dat als je één keer eens verkeerd in klopt en dat het dan constant overgenomen wordt. Dan moet daar eigenlijk een soort tegenhanger voor zijn dat er ergens nog een keer een checkmoment zit van: klopt dit nu nog? We zien het wel eens vaker dat er ergens een in rapportage één gezegd wordt, dat iemand van die en die datum in kliniek x heeft gezeten en dat wordt geciteerd op geciteerd, op geciteerd en dat wordt nooit meer gecheckt. Vijf jaar later blijkt als een cliënt daar nog eens kritisch naar kijkt, dat die dan constateert van hé maar daar heb ik niet gezeten. Maar dat gaat dan wel een eigen leven leiden en dat is moeilijk om dan weer weg te krijgen uit het systeem." (reclasseringswerker 2)*

---

Al met al zijn de reclasseringswerkers tevreden met de RISC als vorm van AI-gebaseerd risicotaxatie-instrument, maar er is geen behoefte om dit verder uit te breiden in termen van AI.

#### 4.3.3.2 AI als informatieclassificator

Alle reclasseringswerkers zien zeker kansen liggen in het gebruik van AI als informatieclassificator. Allereerst wordt als meest belangrijke vorm hiervan gezien dat AI zou kunnen ondersteunen in het vergaren en verwerken van informatie binnen een casus.

---

*"Ik denk wel dan puur in de uitvoering dan he. De meeste casussen die zijn allemaal niet zo heel ingewikkeld en daar kun je zo doorheen werken. Maar soms loop ik tegen onderwerp waarvan je denkt, ja, ik weet ook niet hoe het zit. En als je dan achtergrondinformatie makkelijk kan vergaren... Ja, graag." (reclasseringswerker 1)*

*"Dat dat zou ik wel helpend vinden. Want ik merk nu in de voorbereiding: soms hebben cliënten een enorme lange staat van dienst bij reclassering en justitie en dan is er een heleboel te lezen en ik vind het vaak een nou een hachelijke onderneming om alles tot me te nemen. Dus ik ga vaak niet verder terug dan vijf jaar omdat het soms gewoon geen doen is. Over sommige mensen zijn 40 rapporten over geschreven rapportages en weet ik van wat. Dus als je zoiets hebt, dat het dan op een bepaalde ordening met een relevantie aan je wordt gepresenteerd. Dat kan me wel helpen zijn in mijn voorbereiding op een dossier." (reclasseringswerker 4)*

*"Het proces-verbaal die we soms aangeleverd krijgen, die kunnen ook fors zijn. Het kost heel veel tijd om ze door te lezen. Heb je weinig tijd dan ga je automatisch al proberen en kijken of je dingen kunt overslaan. Als AI alles kan lezen in een enorme snelle, korte tijd, zeg maar, en daar alle relevante informatie kan ordenen. Ja, dan kan ik me voorstellen dat het heel veel tijds winst oplevert. [...] En als ze dan in die risicotaxatie instrumenten ook nog een goeie ontwikkeling kan komen waarbij het heel snel kan. En toch wel goeie informatie er in inzet, ja, dat zou natuurlijk mooi zijn.." (reclasseringswerker 6)*

---

---

---

Hierbij gaat het dus om het gebruik van AI met als doel om relevante informatie uit een dossier te krijgen en deze te presenteren aan de reclasseringswerker. Volgens twee reclasseringswerkers gebeurt het nog vaak dat informatie uit dossiers wordt gekopieerd – of deze nou volledig is of niet – waardoor de rapporten veel dubbele informatie bevat. Een goed AI-systeem zou dit kunnen markeren en aangeven waarin dezelfde informatie wordt benoemd zodat de dubbel gearchiveerde informatie uit het dossier van de cliënt kan worden gehaald.

AI zou ook ingezet moeten worden om informatie te ordenen buiten casussen om, zoals binnen het eigen intranet. Reclasseringswerker 2 merkt hierover op:

---

*"Wat ik niet heel goed vind, dat is onze eigen intranet en onze kennisbank waarin we informatie terug kunnen vinden. Heel vaak ben ik aan het zoeken en dan kan ik iets niet vinden en hij is onlangs verbeterd. Dus het is al wel beter geworden, maar ik denk dat daar echt dat soort dingen dat je echt makkelijk je informatie terug kan vinden. [...] Dat ik het met alle zoektermen die ik kan bedenken niet terug kan vinden en dat ik denk dat niet alle informatie erin staat, maar ook niet altijd wat voor mij dan logische zoektermen zijn leiden tot resultaat."*

---

Met AI zou het mogelijk kunnen zijn om kennisbanken en intranet beter up-to-date te houden en sneller informatie aan elkaar te koppelen. Maar dan moet het AI-programma wel accurate suggesties kunnen doen en voorspellend kunnen werken:

---

*"Ja, dat probleem is natuurlijk dat als je gaat zoeken, dan denk je al dat er iets is. En op moment dat je niet je bewust bent dat er iets over bekend is, ga je ook niet zoeken, dus dat is niet hoe het helemaal zou werken. [...] Je weet niet altijd wat je niet weet."* (reclasseringswerker 3)

*"Ik merk namelijk dat we zoveel informatie hebben en je wordt overspoeld en ja, je kan het vaak niet vinden, het is net niet op die zoektermen die jij dan hebt bedacht en dan moet je nog zestien keer klikken voordat je die paar documenten vindt. Het is altijd één grote wirwar. Dus alleen dat al, qua het toegankelijker maken van informatie en inderdaad slimmer omgaan met het geven van informatie."* (reclasseringswerker 5)

---

AI als informatieclassificator is iets waar geen van de reclasseringswerkers voor het interview kansen in zagen, maar bij het uitvragen hiervan zijn ze allen er over eens dat – wanneer het systeem goed functioneert – dit een belangrijke functie kan zijn om het werk als reclasseringswerker efficiënter uit te kunnen voeren.

#### 4.3.3.3 AI als ondersteunend contact voor de client

Het gebruik van AI in het contact met de cliënt komt futuristisch over. Een voorbeeld van een toepassing is chatbots – bekend van onder andere webwinkels – waaraan een cliënt een vraag zou kunnen stellen aan een computer, het AI-systeem. Hier was van te voren geen wens van reclasseringswerkers om zo iets ook te hebben en vijf van de zes reclasseringswerkers zien dit ook niet zitten en zijn kritisch over deze mogelijke AI-toepassing.

---

*"Nou, daar heb ik heel veel aarzelingen bij en vooral waar het meer gaat om, denk ik, de vragen die wat persoonlijker gericht zijn. Kijk, als dat wat algemener is om te weten van he, hoe werkt, hoe werkt het daarmee, wanneer kan ik dit verwachten? Dat soort algemeenheden dat zou ik niet zoveel moeite mee hebben, geloof ik."* (reclasseringswerker 3)

*"Ik vind het altijd ontzettend irritant als ik met een bedrijf in contact moet komen en dan zit ik tegen een computer te praten. Ik heb het meteen door dat het een chatbot is. Ik haak eigenlijk vrij snel af, ook bij PostNL is dat ook. Ik ga nog liever tien minuten in de wacht staan dan dat ik via zo'n chatbot... Ik vraag me af of ons werk zich er ook voor leent. Ik zou er niet zo goed weten wat het zou kunnen toevoegen. We hebben een bureaudienst die kunnen ze gewoon bellen."* (reclasseringswerker 4)

*"Onze doelgroep, weet je wel, die zijn meestal niet heel handig met telefoons en die hebben meestal niet eens een smartphone.. Maar ook weet je de frustratie dat je dan bijvoorbeeld iets invult en dat je één of ander antwoord krijgt waar je echt helemaal niet zit te wachten, ben ik benieuwd of dan vervolgens heel veel scheldwoorden gaan volgen van deze cliënt."* (reclasseringswerker 5)

---

Eén van de reclasseringswerkers geeft aan dat het wel zou kunnen, maar onder voorwaarden:

---

*"Alleen moeten dan, denk ik, uiteindelijk ook nog wel weer een doorlinkje zitten naar wat je dan nu ook hebt, bij webshops van wil je contact met de medewerker?" (reclasseringswerker 2)*

---

Het moet hierbij voor de cliënt duidelijk zijn dat diegene met een chatbot praten en dat er alleen simpele vragen gesteld kunnen worden.

Zoals reclasseringswerker 6 aangeeft, is het voor een AI-systeem lastig om de toestand van een cliënt in te schatten:

---

*"Daarmee zie je al heel snel dat een ervaren receptioniste, althans in de oude situatie bij ons, heel snel door heeft dat iemand wel of niet goed in z'n vel zit of erg opgefokt is. Bijvoorbeeld, iemand die heel boos is om wat voor reden dan ook en daarmee al heel snel kan schakelen met de reclasseringswerker die naar beneden komt om de klant te spreken of dat je daar een afspraak mee heeft. Nou, dat zijn natuurlijk wel hele lastige dingen, lijkt mij, als je kijkt naar AI." (reclasseringswerker 6)*

---

Naast de chatbots werd geopperd door reclasseringswerker 1 om AI in te zetten bij de huidige e-health modules als het gaat om AI-toepassingen in omgang met cliënten.

---

*"Ja, dat wordt in de psychologie al op ruimschoots toegepast. Dat je dat gewoon echt AI-behandeltrajecten hebt. Ik zal niet weten waarom het binnen het reclasseringswerk niet zou kunnen. Het is ook vaak zo dat mensen psychische klachten hebben als ze met ons in aanraking komen. Die hebben ze ook al voordat ze met ons in aanraking komen. Dus waarom dat niet zo één op één te gebruiken? Ja, dat kan. In principe verwijzen we ook veel cliënten door naar psychische hulp, psychiatrische hulp, al dan niet forensisch. Ik heb nog niet gehoord dat die techniek ook in die trajecten gebruikt wordt. Maar waarom niet? En voor echt in het reclasseringscontact? Ja, ik zou me kunnen voorstellen dat iemand een videodagboek bijhoudt of zo of in ieder geval dat er een dagboek bijgehouden wordt en dat dan ook de AI iemands gezichtsuitdrukking kan herkennen. Van ja, je zegt wel dat het allemaal goed gaat, maar je kop staat op storm." (reclasseringswerker 1)*

---

De toepassing van AI in het contact met de cliënt lijkt momenteel te onduidelijk wat het kan inhouden om op dit moment effectief ingezet te worden binnen 3RO. Verdere ontwikkeling van dit soort AI-programma's en praktische toepassingen moet dus gebeuren.

#### 4.3.3.4 Toepassingen die door reclasseringswerkers zelf werden aangedragen

Naast de door ons geopperde toepassingen van AI zoals hierboven beschreven, kwam reclasseringswerker 1 op het idee om AI toe te passen in het gebruik van de trainingen die reclasseringswerkers moeten volgen. Momenteel wordt er in deze trainingen door reclasseringswerkers getraind in situaties met een acteur. In Virtual Reality (VR) omgevingen is het nu mogelijk om trainingen op te zetten voor een grote variatie aan situaties zonder dat daar een trainingsdag voor opgezet hoeft te worden waar meerdere reclasseringswerkers fysiek aanwezig moeten zijn.

---

*"We merken gewoon en dat is echt een algemeen 3RO, volgens mij, probleem wat ik al ken sinds dat ik er werk, is dat het altijd een onwijs gedoe is om groepen vol te krijgen voor trainingen, voor nou alle mooie methodieken die we hebben aan stap voor stap, apps worden nooit gebruikt." (reclasseringswerker 5)*

*"En dat zal wel wennen zijn, want nu komt er een gewoon een acteur en dat. Maar wat ik begrijp van collega's die daar zich wat meer mee bezig houden, is dat dit ook heel goed helpend is om echt goed real-life te trainen, terwijl je geen acteur nodig hebt." (reclasseringswerker 2)*

---

AI zou in deze VR-omgevingen bij kunnen dragen door op te sporen welk onderdeel van de training nog niet optimaal wordt uitgevoerd door de reclasseringswerker en om vervolgens die specifieke onderdelen herhaaldelijk aan te bieden. Echter, ook hier zit weer een kanttekening aan:

---

*"Het programma moet goed kunnen inschatten dat het niet te ver gaat. Want zoals bijvoorbeeld met zo'n acteur: aan het begin van de training, wordt iedereen vraag van: heb jij ervaringen waar wij rekening mee moeten houden. De acteur die voelt en die ziet wanneer die te ver gaat? Ja, of het programma zo sensitief is, dat weet ik niet." (reclasseringswerker 1)*

---

Een andere toepassing van AI is om meer met de data te doen die middels de elektronische monitoring wordt verkregen, zoals aangegeven door reclasseringswerker 3 en 5:

---

*"Alleen ik denk dat het gebruik beperkter is voor alles waar ze het in zouden willen toepassen. Kijk op het moment dat je van belang vindt dat iemand niet in de buurt van scholen komt en je hebt een gebied of een locatiegebod of iets dergelijks. Dat zegt nog steeds niet dat iemand niet in contact komt met. Dus dan moet je heel erg bewust zijn van inderdaad een hulpmiddel is en niet een middel wat voorkomt dat er een recidief plaatsvindt of dat er toch zaken gebeuren die volgens de regels niet zouden mogen. Dus dat is het beperkte bruikbaarheid van het middel." (reclasseringswerker 3)*

*"Maar goed, dat is natuurlijk makkelijk als een computer of een systeem kan koppelen van hé, maar die locatie is een basisschool, of die locatie is puntje, puntje, puntje, dat zou echt wel... Dat scheelt je ook als werker heel veel zoekwerk want anders moet je de hele dag moet je bijna die trillingen gaan zitten uitlezen waar iemand dan precies was en dan moet je met Streetview gaan kijken wat dat dan is. Dat zou natuurlijk qua tijdsinstaat zou dat wel heel mooi zijn, ja." (reclasseringswerker 5)*

---

Het zou dus kunnen gaan om het gesprek aan te gaan met een cliënt, waarom hij of zij vaak tijd doorbrengt in bepaalde gebieden en hoe hem of haar daarmee verder te helpen.

Echter, deze toepassing moet wel nauwkeurig gebruikt worden. Allereerst in wie er toegang heeft tot de data.

---

*"Als je iedereen in dat systeem laat. Dan heb je d'r weinig controle meer op wat mensen wel en niet zien en dan wordt het een beetje een soort geleidende schaal naar ja, dingen waar wij dat daarin niet zeg maar echt verantwoord vinden, ethisch gezien ook" (reclasseringswerker 5)*

---

Daarnaast moet het wel effectief ingezet kunnen worden.

---

*"Kijk, als je de elektronische monitoring heeft voor mijn gevoel altijd een beetje de neiging dat je meer aan de kant van de politie en aan de controlerende kant zit waar je niet meer de toegevoegde waarde hebt waar je als reclassering hybride werkt. Dat zijn van die grensgebieden waarvan ik vind dat je erg moet waken hoe ze in te zetten en hoe ze te gebruiken omdat je niet de controleur bent" (reclasseringswerker 3)*

---

De reclasseringswerkers hebben dus zelf ook goede ideeën over de toepassing van AI binnen 3RO wanneer ze ernaar gevraagd worden. Echter, door de ontwikkeling van de programma's is het voor hen vaak nog niet duidelijk of zulke ideeën ook haalbaar zijn.

#### 4.3.4 Stappen nodig om AI binnen 3RO succesvol te implementeren

Elk van de reclasseringswerkers heeft een eigen visie op wanneer AI binnen 3RO succesvol is geïmplementeerd.

---

*"Dat het helder en duidelijk is wat het doet, waarom het dat doet, wat je zelf moet doen." (reclasseringswerker 1)*

*"Ik denk dat het succesvol is toegepast als het de reclasseringswerker ontzorgt. Dat kan praktisch zijn van dat het ergens tijd bespaart en iets. Of dat het zorgt dat de kwaliteit van het werk groter wordt. Dan is het geslaagd. Ja, volgens mij is dat dat." (reclasseringswerker 2)*

---

---

*"Ja, dat enerzijds denk ik dat het wat ik zei, dat besef hé, dat algoritmes naast je eigen professioneel oordeel of in combinatie met, dat die bevindingen ergens op een goeie manier verankerd zouden moeten worden. En hoe je dat moet doen dat is, ja, dat is gewoon heel ingewikkeld." (reclasseringswerker 3)*

*"Dat ik nog wel de ruimte krijg om mijn ervaringen met die cliënt ook... Want die computer spreekt die cliënt niet, want dat doe ik. En dat ik dat daar wel nog in kwijt kan en dat ik ook de vrijheid heb om... Het moet niet leidend zijn, vind ik dat hele systeem met het AI gebeuren. Dus ik wil daar nog wel van kunnen afwijken." (reclasseringswerker 4)*

*"Dus ik denk wel dat je mensen erg mee moet nemen in de voordelen ervan. En ook hé, dat ze niet het gevoel hebben dat oh nou wordt mijn werk overgenomen, of wat ik vind is niet-relevant of een computer gaat even vertellen hoe het moet." (reclasseringswerker 5)*

*"Maar ja, maar als je kijkt naar de toekomst, wanneer is het dan succesvol? Ja, als het echt bijdraagt aan gedragsbeïnvloeding in positieve zin van onze reclasseringscliënten en daarmee ook de herhaling van strafbare feiten, recidive vermindering plaatsvindt. Dan is het wat mij betreft succesvol." (reclasseringswerker 6)*

---

AI-programma's moeten dus kunnen bijdragen aan efficiënter werken met een positief effect op de cliënten. Er moet daarbij wel altijd de ruimte blijven voor de inbreng van de reclasseringswerker zelf: alle reclasseringswerkers zijn het er over eens dat beslissingen niet overgenomen moeten worden door welk AI-programma dan ook.

---

*"Nee, het belangrijkste voor mij is dat wij reclasseringswerkers zijn en dat wij een toegevoegde waarde hebben als reclasseringswerker en dat is niet door middel van alleen het toepassen van instrumenten." (reclasseringswerker 3)*

*"Het moet niet leidend zijn, vind ik dat hele systeem met het AI gebeuren. Dus ik wil daar nog wel van kunnen afwijken." (reclasseringswerker 4)*

---

Of AI-programma's – in welke vorm dan ook – geaccepteerd zullen worden, is lastig te overzien voor de reclasseringswerkers:

---

*"Wat met name gewoon echt wel iets is van reclasseringswerkers wat ik al aan het begin zei, is dat we gewoon niet zo heel erg vertrouwen hebben in de computergestuurde programma's die jou vertellen hoe je iets moet doen of zo. Dus ik denk wel dat je mensen erg mee moet nemen in de voordelen ervan." (reclasseringswerker 5)*

*"Ja, dat enerzijds denk ik dat het wat ik zei, dat besef hé, dat algoritmes naast je eigen professioneel oordeel of in combinatie met, dat die bevindingen ergens op een goeie manier verankerd zouden moeten worden." (reclasseringswerker 3)*

---

Eén reclasseringswerker merkt op dat de term AI ook in het Nederlands kan worden gebruikt: kunstmatige intelligentie. Dit zou bevorderlijk kunnen zijn voor acceptatie van AI programma's binnen 3RO.

---

*"Ja, en ik denk op zich dat als je het letterlijk naar Nederlands vertaalt, de kunstmatige intelligentie, dat die dan al iets tastbaarder wordt. In ieder geval misschien voor mijn generatie, die dan daar meer beeld bij heeft als bij Artificial Intelligence. Dan denk ik aan iets abstracter, iets groters wat verder weg staat." (reclasseringswerker 2).*

---

Ook zullen de reclasseringswerkers zelf betrokken moeten zijn bij het ontwikkelen van zulke programma's: *"Ik denk daar waar je input moet hebben over praktijksituaties dus dat je daar wel mensen van het 3RO moet bij betrekken." (reclasseringswerker 2).*

#### 4.3.5 Conclusie interviews met reclasseringswerkers

Uit de interviews met de reclasseringswerkers blijkt dat alhoewel er enige basiskennis is van AI, de mogelijkheden die AI biedt voor reclasseringswerkers binnen 3RO niet altijd helder zijn (onderzoeksvraag 1 en 2). Er zijn vele mogelijkheden buiten de risicotaxatie-instrumenten die op basis van AI werken, maar toch zijn AI-toepassingen in risicotaxatie-instrumenten de eerste gedachte van de reclasseringswerkers (onderzoeksvraag 2 en 3). Het is helder geworden dat er geen

behoefte is om deze risicotaxatie-instrumenten verder uit te breiden, maar om met de RISC verder te gaan in de huidige vorm. Qua gebruiksvriendelijkheid laat deze nog iets te wensen over, wat een goede verbetering zou zijn.

Eenieder is enthousiast over het gebruik van AI als informatieclassificator. Alhoewel niemand hier van te voren wensen over had, is het duidelijk dat een snelle ordening van informatie binnen dossiers van cliënten en eventueel een goed werkend intranet/kennisbank een goede bijdrage kan leveren aan het werk van de reclasseringswerker (onderzoeksvraag 4, 5 en 6).

Chatbots en gelijke AI-toepassingen zijn in een beginstadium van hun ontwikkeling. Het gebrek aan mogelijkheden om complexe vragen te beantwoorden van cliënten en eventuele problemen in gebruik door cliënten die minder digitaal vaardig zijn, zorgt ervoor dat dit soort AI-toepassingen veel verder door ontwikkeld moeten worden voordat 3RO hier mee kan beginnen te werken.

AI binnen 3RO kan zijn meerwaarde hebben (onderzoeksvraag 7 en 8). Het belangrijkste is dat AI vooral moet kunnen ontzorgen en ontlasten voor de reclasseringswerkers, ongeacht wat de exacte AI-toepassing nou precies is. Als duidelijk is hoe AI-programma's dit kunnen doen, zal er ook snellere acceptatie van de programma's door de reclasseringswerkers zijn. Echter, het staat vast dat de beslissingen over cliënten bij de reclasseringswerker moeten blijven liggen. Zoals reclasseringswerker 4 het zegt:

---

*"Uiteindelijk is mijn werk mensenwerk en niet computerwerk."* (reclasseringswerker 4)

---

## 5. Discussie

Het doel van dit rapport was om voor het reclasseringswerk beter te begrijpen wat AI behelst, hoe AI ingezet kan worden binnen het reclasseringswerk, wat de voor- en nadelen van deze AI-toepassingen zijn en hoe reclasseringswerkers met de toevoeging van zulke AI-programma's om kunnen gaan.

Om dit doel te bereiken hebben wij een scoping review uitgevoerd en interviews afgenomen met experts en met reclasseringswerkers. De volgende onderzoeksvragen stonden centraal in dit onderzoek.

1. Wat is de huidige stand van zaken in AI in een justitiële en GGZ-context, en wat is de meest passende definitie van AI?
2. Hoe wordt AI gezien door medewerkers binnen de 3RO?
3. Waar en hoe wordt AI al binnen de context van het reclasseringswerk toegepast?
4. Wat zijn mogelijke toepassingsmogelijkheden van AI binnen het reclasseringswerk?
5. Wat zijn mogelijke voor- en nadelen van de verschillende toepassingsmogelijkheden van AI in het reclasseringswerk?
6. Wat zijn morele, ethische en juridische aspecten die een rol spelen in de beslissing om AI in het reclasseringswerk toe te passen?
7. Welke vervolgstappen zijn er nodig in onderzoek naar AI binnen het reclasseringswerk?
8. Welke vervolgstappen zijn er nodig in de praktijk om AI goed toe te passen binnen het reclasseringswerk?

Op basis van de bevindingen uit de scoping review en de interviews kunnen we deze onderzoeksvragen beantwoorden.

### 5.1 Huidige toepassing AI binnen justitiële veld en 3RO (onderzoeksvraag 1 en 3)

Het is duidelijk dat AI binnen het justitiële veld in opkomst is. Steeds meer AI-programma's worden ontwikkeld en in gebruik genomen. In Nederland gaat dit momenteel voornamelijk om het gebruik van predictive policing in het justitiële veld. Internationaal gezien betreft het ook het gebruik van het recidive-instrument COMPAS. Binnen 3RO specifiek is het risicotaxatie-instrument de RISC (met daarin de OXREC) momenteel het programma waarin AI-elementen verwerkt zitten. Vanuit de interviews met reclasseringswerker blijkt ook dat binnen Europese reclasseringsorganisaties er behoefte is om te onderzoeken wat technologie in hun organisatie kan bijdragen, waarbij AI een prominent onderdeel is van de mogelijk bruikbare technologie.

Vanuit voornamelijk de scoping review wordt duidelijk dat binnen het justitiële veld vooral aan risicotaxatie-instrumenten gedacht wordt als het gaat om AI-toepassingen. AI als informatieclassificatie wordt nog weinig toegepast. Echter – vanuit voornamelijk de interviews met reclasseringswerkers – wordt duidelijk dat er geen behoefte is aan verdere ontwikkeling van de AI-component in de risicotaxatie-instrumenten, maar dat iedereen enthousiast wordt van het toepassen van AI om een beter intranet te krijgen of een betere zoekfunctie in een database. AI kan in dat opzicht dus veel bieden voor het reclasseringswerk.

Geconstateerd is een gebrek aan kennis over wat AI zou kunnen betekenen voor het reclasseringswerk. Dit hangt mogelijk ook samen met het ontbreken van een goede definitie over wat AI precies inhoudt. Zowel in de scoping review als in de interviews met experts komt hier geen eenduidig antwoord naar voren. Het gaat om het ontwikkelen van intelligente systemen, maar wat deze intelligente systemen precies zijn, wat zij kunnen doen en of ze evenveel of meer kunnen bereiken dan de mens blijft zich verschuiven in de loop der tijd. Een eenduidige definitie van AI zou mogelijk ook kunnen bijdragen aan de ontwikkeling van AI-systemen.

### 5.2 AI door de ogen van de reclasseringswerker (onderzoeksvraag 2)

Bij de reclasseringswerkers is er al kennis van AI in het dagelijks leven. Zoals al gezegd was er echter minder kennis over de toepassing van AI in het reclasseringswerk buiten de RISC en andere risicotaxatie-instrumenten om. In de huidige AI-toepassingen of AI-toepassingen die nog ontwikkeld moeten worden, zijn alle reclasseringswerkers het er over eens dat de menselijke maat in het reclasseringswerk moet blijven. Dit zonder meer over laten aan een AI-systeem is niet wenselijk: de beslissingen die genomen worden hebben een grote impact op het leven van een cliënt. Alhoewel het niet waarschijnlijk is dat dit ooit zal gebeuren (Morison & Harkens, 2019) is er hier duidelijk wel



enige angst voor vanuit de positie als reclasseringswerker. Acceptatie van AI-systemen binnen 3RO door de reclasseringswerkers zal dus alleen mogelijk zijn zodra de systemen een duidelijk toegevoegde waarde hebben zonder de autonomie van de reclasseringswerker aan te tasten.

### 5.3 De toekomst in toepassing van AI binnen 3RO (onderzoeksvraag 4 tot en met 8)

Wanneer AI-systemen goed functioneren kan dit er voor zorgen dat er potentieel werk uit de handen wordt genomen van de reclasseringswerkers en dit kan bijdragen aan het efficiënter werken binnen 3RO. Er liggen kansen voor 3RO om AI-systemen te ontwikkelen die de reclasseringswerkers kunnen ontzorgen in hun werkzaamheden, met name bij het ordenen en aanbieden van relevante informatie. Dit kan door zoals eerder aangegeven een handig en overzichtelijke intranet en databases te ontwikkelen waarin de crux zit dat deze systemen niet simpelweg zoekmachines zijn, maar ook goed kunnen voorspellen waar behoefte aan is middels de ingevoerde zoekopdrachten door de reclasseringswerkers.

Het is nog niet optimaal om AI-systemen in te zetten in het contact met de cliënt. Simpele vragen zouden beantwoord kunnen worden door een chatbot, maar complexe vragen lijken AI-systemen (nog) niet aan te kunnen. Hierbij lijkt het AI-systeem dan niet beter te kunnen functioneren dan de huidige middelen online of de hulpdienst die gebeld kan worden. Vanuit de interviews met reclasseringswerkers lijkt dat het er wel op dat binnen de reclassering van de Verenigde Staten onderzocht wordt hoe cliëntmotivatie verhoogd kan worden door AI-systemen motiverende teksten te personaliseren naar de doelen van de cliënt. Echter, dit onderzoek is nog niet openbaar en lijkt zich nog in een vroege fase te bevinden. De digitale vaardigheden van de cliënten om met zo'n AI-systeem om te gaan en de soms beperkte cognitieve vaardigheden om te beseffen dat er met een AI-systeem contact is en niet met een mens, maakt dat dit cliënt contact door middel van AI voorlopig nog toekomstmuziek is.

Echter, duidelijk is dat er nog veel haken en ogen zitten aan de inzet van AI-systemen. Dit zit hem voornamelijk in de mogelijke bias, het black-box effect en de acceptatie van AI-systemen door gebruikers. Voornamelijk de aanwezigheid van de bias die in de beslissingen van AI risicotaxatie-instrumenten aanwezig is – en altijd in enige mate aanwezig zal blijven – maakt dat de verdere inzet van AI in risicotaxatie-instrumenten binnen 3RO goed overwogen moet worden. Er lijkt namelijk geen verdere behoefte vanuit reclasseringswerkers om het AI-component in deze instrumenten uit te breiden.

Een limitatie van het huidige onderzoek is het gebrek aan contact met cliënten en hun mening over AI. AI-systemen (en dan voornamelijk risicotaxatie-instrumenten) kunnen een grote impact hebben op hun leven. Als er concreet naar de RISC wordt gekeken, is het niet duidelijk hoe cliënten omgaan met de score vanuit de OXREC en de (mogelijk) afwijkende beslissing van de reclasseringswerker(s). Cliënten verwachten een eerlijk en onbevooroordeelde beslissing. Hun acceptatie van deze beslissingen zal beïnvloed kunnen worden door het gebruik van AI-systemen (is het systeem wel optimaal genoeg om deze individu te beoordelen?) of juist de afwezigheid van deze systemen (is de reclasseringswerker niet bevooroordeeld?).

### 5.4 Concrete aanbevelingen

Op basis van ons onderzoek willen wij graag een aantal concrete aanbevelingen doen als het gaat om het verder ontwikkelen van AI in 3RO.

#### 5.4.1 Voorbij AI als risico-taxatie

Uit ons onderzoek blijkt dat er mogelijkheden zijn om AI toe te passen binnen 3RO waar misschien niet gelijk aan gedacht wordt. Er blijkt geen nadere behoefte om verdere risico-taxatie instrumenten te ontwikkelen op basis van AI, maar juist in AI-systemen om het huidige werk nog beter uit te kunnen voeren. Dit gaat dan vooral om AI als informatieclassificator. Een goed werkend intranetsysteem op basis van AI waarbij het systeem de juiste informatie snel kan vinden en zelfs relevante informatie kan aanbieden waar nog niet aan is gedacht, kan leiden tot meer efficiëntie. Ook AI in cliëntendossiers kan bijdragen aan een lagere werkdruk, door bijvoorbeeld (zoals gesuggereerd door reclasseringswerkers) aan te geven welke informatie in de vele dossiers dubbel is en niet extra hoeft worden doorgelezen. De precieze details hoe zo'n systeem kan werken zal nog nader uitgewerkt moeten worden binnen 3RO, maar duidelijk is dat er behoefte is aan dit soort systemen. AI-systemen samen ontwikkelen met de werkvloer zal zorgen voor een groter draagvlak en acceptatie voor deze systemen.

#### 5.4.2 Kijk naar de ontwikkelingen binnen (explainable) AI

Er is al veel mogelijk binnen AI en er zal waarschijnlijk nog veel meer mogelijk worden in de toekomst. Het is belangrijk om deze ontwikkelingen in de gaten te houden: zowel de positieve als negatieve ontwikkelingen. Het ontwikkelen van AI-systemen kan kostbaar zijn in tijd, inzet en financiën. Sommige systemen (zoals risicotaxatie-instrumenten op basis van AI) zijn al klaar om toegepast te worden terwijl andere systemen (zoals explainable AI waarbij AI-systemen die beslissingen kunnen nemen die voorheen enkel alleen door mensen konden worden uitgevoerd en kan uitleggen waarom deze beslissing wordt gemaakt) nog volop in ontwikkeling zijn. Er moet goed in de gaten gehouden worden waar kansen liggen voor 3RO, wanneer het moment is om in te stappen in een AI-systeem en vooral ook met welke partners (particulier, universiteit of zelf-ontwikkelde AI). Dit is een voorwaarde om verder te kunnen gaan met AI binnen 3RO. Bestuurders binnen 3RO moeten ondervinden en uitzoeken in welke mate zij controle willen hebben in de ontwikkeling van een AI-systeem en of het ten opzichte van mogelijke problemen rondom privacy, tijd en financiën beter is om dit geheel zelf te laten ontwikkelen voor 3RO.

#### 5.4.3 Verlies de menselijke maat niet

Ongeacht waar en hoe AI binnen 3RO ingezet wordt of gaat worden is de boodschap duidelijk: verlies de menselijke maat niet. 3RO opereert met een bijzondere groep cliënten en neemt beslissingen die levens kunnen veranderen. Als AI-systemen deze beslissingen nemen of informatie aanlevert die deze beslissingen kunnen beïnvloeden, is het van belang om de reclasseringswerker het eindoordeel te gunnen. Ook moet deze menselijke maat in de ontwikkeling van AI-systemen worden meegenomen.

### 5.5 Conclusie

De mogelijkheden van AI zijn veelvuldig en wat er nog mogelijk gaat worden kunnen wij nog niet bevatten. Voor wat betreft het toepassen van AI binnen het justitiële veld en binnen 3RO in het bijzonder zitten we nog in de beginfase. Er is nog veel onduidelijk over wanneer we AI het beste kunnen inzetten, wanneer AI het niet bij het rechte eind heeft en wanneer het menselijke oordeel het meest wenselijk is. We zullen kritisch moeten blijven op AI-systemen en deze systemen continue moeten monitoren in hun inzet met oog voor de menselijke maat.

## 6. Referenties

- Alikhademi, K., Drobina, E., & Prioleau, D. (2022). A review of predictive policing from the perspective of fairness. *Artificial Intelligence and Law*, 30(1), 1–17. <https://doi.org/10.1007/s10506-021-09286-4>
- Bagaric, M., Hunter, D., & Stobbs, N. (2019). Erasing the bias against using artificial intelligence to predict future criminality: algorithms are color blind and never tire. *University Clinical Law Review*, 88, 1037–1082.
- Beretta, E., Santangelo, A., Lepri, B., Vetrò, A., & De Martin, J. C. (2019). The Invisible Power of Fairness. How Machine Learning Shapes Democracy. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11489 LNAI, 238–250. [https://doi.org/10.1007/978-3-030-18305-9\\_19](https://doi.org/10.1007/978-3-030-18305-9_19)
- Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology*, 13, 193–216. <https://doi.org/10.1007/s11292-017-9286-2>
- Bicakci, Y. S., Seker, D. Z., & Demirel, H. (2020). Location-based analyses for electronic monitoring of parolees. *ISPRS International Journal of Geo-Information*, 9(5). <https://doi.org/10.3390/ijgi9050296>
- Boahemaa, F. J. (2019). The impact of artificial intelligence on justice systems. In *Trento Biolaw Selected Student papers*. <https://doi.org/10.5958/0976-5506.2019.01876.X>
- Brayne, S. (2020). *Technologies of Crime Prediction : The Reception of Algorithms in Policing and Criminal Courts*. 512, 1–17. <https://doi.org/10.1093/socpro/spaa004>
- Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the Predictive Validity of the Compas Risk and Needs Assessment System. *Criminal Justice and Behavior*, 36(1), 21–40. <https://doi.org/10.1177/0093854808326545>
- Browning, M. (2021). Stop and Risk : Policing , data, and the digital age of discrimination. *American Journal of Criminal Justice*, 46, 298–316.
- Carlson, A. M. (2017). The need for transparency in the age of predictive sentencing algorithms. *Iowa Law Review*, 103(1), 303–329.
- Csányi, G. M., Nagy, D., Vági, R., Vadász, J. P., & Orosz, T. (2021). Challenges and open problems of legal document anonymization. *Symmetry*, 13(8), 1–25. <https://doi.org/10.3390/sym13081490>
- Deeks, A. S. (2018). Predicting enemies. *Virginia Law Review*, 104(8), 1529–1592.
- Demura, M., & Klepka, D. (2021). Using artificial intelligence algorithms in the field of criminal judiciary: International experience and domestic prospects. *Scientifi Innovation*, 17(5), 95–101.
- Docarmo, T., Emery, J., Rea, S., & Raval, N. (2021). The law in computation : What machine learning , artificial intelligence , and big data mean for law and society scholarship. *Law & Policy*, 43, 170–199. <https://doi.org/10.1111/lapo.12164>
- Dressel, J., & Farid, H. (2018). The accuracy , fairness , and limits of predicting recidivism. *Science Advances*, March, 1–6.
- Duker, M. J. A., & Lodder, A. R. (1999). Sentencing and information management: Consistency and the particularities of a case. *Proceedings of the International Conference on Artificial Intelligence and Law*, c, 100–107.
- Eckhouse, L., Lum, K., Conti-Cook, C., & Ciccolini, J. (2019). Layers of Bias: A Unified Approach for Understanding Problems With Risk Assessment. *Criminal Justice and Behavior*, 46(2), 185–209. <https://doi.org/10.1177/0093854818811379>
- Erdem, G., Tuncer, A. E., Safi, O. A., Çankaya, B., Ergin, M., & Aydoğan, R. (2019). The professional experiences and training needs of probation officers in Turkey. *Journal of Social Work*, 19(5), 664–686. <https://doi.org/10.1177/1468017318768181>
- Fortes, P. R. B. (2020). Paths to Digital Justice: Judicial Robots, Algorithmic Decision-Making, and Due Process. *Asian Journal of Law and Society*, 7(3), 453–469. <https://doi.org/10.1017/als.2020.12>
- Franssen, V., & Berrendorf, A. (2021). The use of AI tools in criminal courts: Justice done and seen to be done? *Revue Internationale de Droit Penal*, 92(1), 199–223.
- Giannini, A. (2021). Lombroso 2.0: On AI and predictions of dangerousness in criminal justice. *RIPD*, 92(1), 179–198. <http://arxiv.org/abs/1611.04135>
- Green, B. (2020). The false promise of risk assessments: Epistemic reform and the limits of fairness. *FAT\* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 594–606. <https://doi.org/10.1145/3351095.3372869>
- Hälterlein, J. (2021). Epistemologies of predictive policing: Mathematical social science, social physics and machine learning. *Big Data and Society*, 8(1). <https://doi.org/10.1177/20539517211003118>
- Harkens, A. (2018). The ghost in the legal machine: algorithmic governmentality, economy, and

- the practice of law. *Journal of Information, Communication and Ethics in Society*, 16(1), 16–31. <https://doi.org/10.1108/JICES-09-2016-0038>
- Hilt, K. (2017). What does the future hold for the law librarian in the advent of artificial intelligence? | Que réserve l'avenir pour le bibliothécaire de droit avec la venue de l'intelligence artificielle? *Canadian Journal of Information and Library Science*, 41(3), 211–227. <https://muse.jhu.edu/article/686190/pdf>
- Lee, Y., Soohyun, O., & Eck, J. E. (2020). A Theory-Driven Algorithm for Real-Time Crime Hot Spot Forecasting. *Police Quarterly*, 23(2), 174–201. <https://doi.org/10.1177/1098611119887809>
- Liu, Y., Luo, X., & Yang, X. (2019). Semantics and Structure Based Recommendation of Similar Legal Cases. *Proceedings of IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering, ISKE 2019*, 388–395. <https://doi.org/10.1109/ISKE47853.2019.9170379>
- McKay, C. (2020). Predicting risk in criminal procedure: actuarial tools, algorithms, AI and judicial decision-making. *Current Issues in Criminal Justice*, 32(1), 22–39. <https://doi.org/10.1080/10345329.2019.1658694>
- Ministry of social and family development Singapore. (2021). *Probation and community rehabilitation service*.
- Morison, J., & Harkens, A. (2019). Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making. *Legal Studies*, 39(4), 618–635. <https://doi.org/10.1017/lst.2019.5>
- Moses, L. B. (2020). Artificial intelligence in the courts, legal academia and legal practice. *Australian Law Journal*, 97, 561–574.
- O'Donnell, R. M. (2022). Challenging racist predictive policing algorithms. *New York University Law Review*, 94(55), 1–34.
- Oudekerk, B., & Kaebler, D. (2021). Probation and Parole in the United States, 2019. In *Bureau of Justice Statistics Bulletin: Vol. July*.
- Peeters, R., & Schuilenburg, M. (2018). Machine justice: Governing security through the bureaucracy of algorithms. *Information Polity*, 23(3), 267–280. <https://doi.org/10.3233/IP-180074>
- Ruhl, J. B., Nay, J., & Gilligan, J. (2018). Topic Modeling the President : Conventional and Computational Methods. *The George Washington Law Review*, 86(5), 1243–1315.
- Scurich, N., & Krauss, D. A. (2020). Public's views of risk assessment algorithms and pretrial decision making. *Psychology, Public Policy, and Law*, 26(1), 1–9. <https://doi.org/10.1037/law0000219>
- Shapiro, A. (2019). Predictive policing for reform? Indeterminacy and intervention in big data policing. *Surveillance and Society*, 17(3–4), 456–472. <https://doi.org/10.24908/ss.v17i3/4.10410>
- Sil, R., & Roy, A. (2019). Artificial Intelligence and Machine Learning based Legal Application : The State-of-the-Art and Future Research Trends. *WAIEL*, 57–62.
- Special Interest Group on Artificial Intelligence. (2018). *Dutch Artificial Intelligence Manifesto* (Vol. 14, Issue August).
- Stanford University. (2016). *One hundred year study on Artificial Intelligence*. <https://ai100.stanford.edu/2016-report/section-i-what-artificialintelligence/%0Adefining-ai>
- Swofford, H., & Champod, C. (2021). Implementation of algorithms in pattern & impression evidence: A responsible and practical roadmap. *Forensic Science International: Synergy*, 3, 100142. <https://doi.org/10.1016/j.fsisyn.2021.100142>
- Ting, M. H., Chu, C. M., Zeng, G., Li, D., & Chng, G. S. (2018). Predicting recidivism among youth offenders: Augmenting professional judgement with machine learning algorithms. *Journal of Social Work*, 18(6), 631–649. <https://doi.org/10.1177/1468017317743137>
- Vaccaro, M., & Waldo, J. (2019). The effects of mixing machine learning & human judgment. *Communication of the ACM*, 62(11), 104–110.
- Van den Berg, C., Bruggeman, M., Houston, R., Joosten, A., & Harte, J. M. (2013). Validatiestudie risico- en beschermende factoren van de RISC: Een evaluatieonderzoek naar de leefgebieden van het risicotaxatie en adviesinstrument van de 3RO. In *Vrije Universiteit, Faculteit der Rechtsgeleerdheid* (Vol. 47, Issue 3). <https://doi.org/10.1177/0095399713509530>
- Van den Berg, C., Bruggeman, M., Houston, R., Joosten, A., & Harte, J. M. (2021). Validatiestudie risico- en beschermende factoren van de RISC. In *Vrije Universiteit, Faculteit der Rechtsgeleerdheid*. <https://doi.org/10.1177/0095399713509530>
- Vincent, G. M., & Viljoen, J. L. (2020). Racist Algorithms or Systemic Problems? Risk Assessments and Racial Disparities. *Criminal Justice and Behavior*, 47(12), 1576–1584. <https://doi.org/10.1177/0093854820954501>
- Vries, M. De, Bijlsma, J., Mackor, A. R., Bex, F., & Meynen, G. (2020). *AI-risicotaxatie : nieuwe kansen en risico 's voor statistische voorspellingen van recidive Met name in de Verenigde Staten worden nieuwe risicotaxatie-instrumenten ontwikkeld die kunstmatige intelligentie ( AI*

- ) *toepassen om recidiverisico te voorspelle*. 1784–1790.
- Werth, R. (2019). Risk and punishment: The recent history and uncertain future of actuarial, algorithmic, and “evidence-based” penal techniques. *Sociology Compass*, 13(2), 1–19. <https://doi.org/10.1111/soc4.12659>
- Wright, S. A. (2020). AI in the Law : Towards Assessing Ethical Risks. *IEEE International Conference on Big Data (Big Data) AI*, 2160–2169.
- Žliobaitė, I., & Custers, B. (2016). Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models. *Artificial Intelligence and Law*, 24(2), 183–201. <https://doi.org/10.1007/s10506-016-9182-5>

## Appendix A: Scoping review

**Tabel 2.**

*Zoekopdrachten per database en het aantal hits (inclusief duplicaten)*

<b>Database</b>	<b>Zoekopdrachten</b>	<b>Hits (inclusief duplicaten)</b>
Scopus	TITLE-ABS-KEY ("artificial* intelligen*" OR "machine learn*" OR "algorithm*") AND TITLE-ABS-KEY ("probation*" OR "parole*" OR "legal judgement*" OR "legal practice*" OR "criminal justice" OR "forensic assessment*" OR "forensic mental health" OR "judicial risk assessment*")	579
Web of Science	TS=("artificial* intelligen*" OR "machine learn*" OR "algorithm*") AND TS=("probation*" OR "parole*" OR "legal judgement*" OR "legal practice*" OR "criminal justice" OR "forensic assessment*" OR "forensic mental health" OR "judicial risk assessment*")	221
ACM DL	Title:("artificial* intelligen*" OR "machine learn*" OR "algorithm*") AND Abstract:("probation*" OR "parole*" OR "legal judgement" OR "legal practice" OR "criminal justice" OR "forensic assessment*" OR "forensic mental health" OR "judicial risk assessment*")	14
PubMed	("artificial* intelligen*" OR "machine learn*" OR "algorithm*") AND ("probation*" OR "parole*" OR "legal judgement*" OR "legal practice*" OR "criminal justice" OR "forensic assessment*" OR "forensic mental health" OR "judicial risk assessment*")	120
PsycInfo	TI-AB-KW ("artificial* intelligen*" OR "machine learn*" OR "algorithm*") AND TI-AB-KW ("probation*" OR "parole*" OR "legal judgement*" OR "legal practice*" OR "criminal justice" OR "forensic assessment*" OR "forensic mental health" OR "judicial risk assessment*")	100

**Tabel 3.**

Overzicht papers geïncludeerd in scoping review

Auteurs en Jaar		Titel	Type onderzoek	Onderzoeksveld
1. Alikhademi, K.	2022	A review of predictive policing from the perspective of fairness	Literatuuronderzoek	Justitie
2. Ashley, K. D.	1992	Case-based reasoning and its implications for legal expert systems	Literatuuronderzoek	Justitie
3. Axpe, M. R. V.	2021	Ethical challenges from artificial intelligence to legal practice	Literatuuronderzoek	Justitie
4. Bagaric, M.	2020	Erasing the bias against using artificial intelligence to predict future criminality: algorithms are color blind and never tire	Literatuuronderzoek	Reclassering
5. Bampasika, L. L. M.	2020	Artificial intelligence as evidence in criminal trial	Literatuuronderzoek	Justitie
6. Beretta	2019	The invisible power of fairness. How machine learning shapes democracy	Review paper	AI
7. Berk, R.	2020	Almost politically acceptable criminal justice risk assessment	Case studie	Justitie
8. Berk, R.	2021	Fairness in criminal justice risk assessments: the state of the art	Literatuuronderzoek	Justitie
9. Berk, R.	2017	An impact assessment of machine learning risk forecasts on parole board decisions and recidivism	Experiment	Reclassering
10. Berk, R.	2019	Accuracy and fairness for juvenile justice risk assessment	Literatuuronderzoek	Justitie
11. Bicakci	2020	Location-Based Analyses for Electronic Monitoring of Parolees	Experiment	Reclassering
12. Boahemaa	2018	The impact of Artificial Intelligence on justice systems	Review paper	Justitie
13. Bosse	2019	Sociale kunstmatige intelligentie	Review paper	AI
14. Brayne	2020	Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts	Case studie	Justitie
15. Brennan, T.	2009	Evaluating the predictive validity of the compas risk and needs assessment system	Experiment	Justitie
16. Browning	2021	Stop and risk: policing, data, and the digital age of discrimination	Review paper	Justitie
17. Carlson	2017	The need for transparency in the age of predictive sentencing algorithms.	Review paper	Justitie
18. Chiao, V.	2019	Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice	Literatuuronderzoek	Justitie

19. Christin, A.	2017	Algorithms in practice: comparing web journalism and criminal justice	Case studies	Justitie
20. Chugh	2021	Risk assessment tools on trial: Lessons learned for "Ethical AI" in the criminal justice system	Review paper	Justitie
21. Csányi, G. M.	2021	Challenges and open problems of legal document anonymization	Case studies	Justitie
22. Das, P.	2019	A graph based clustering approach for relation extraction from crime data	Experiment	Justitie
23. De Vries, M.	2021	AI-risicotaxatie: nieuwe kansen en risico's voor statistische voorspellingen van recidive	Review paper	Justitie
24. Deeks	2018	The judicial demand for explainable artificial intelligence	Review paper	Justitie
25. Deeks	2018	Predicting enemies	Review paper	Justitie
26. Delgado	2019	Machine learning in legal practice	Review paper	Justitie
27. Demura	2021	Using artificial intelligence algorithms in the field of criminal judiciary: International experience and domestic prospects	Review paper	Justitie
28. DoCarmo	2021	The law in computation: what machine learning, artificial intelligence, and big data mean for law and society scholarship	Review paper	Justitie
29. Dominic, S.	2016	A review of a face detection system	Literatuuronderzoek	Justitie
30. Dressel	2018	The accuracy, fairness and limits of predicting recidivism	Experiment	Justitie
31. du Plessis, T.	2006	Knowledge management and legal practice	Literatuuronderzoek	AI
32. Duker	1999	Sentencing and Information Management: Consistency and the Particularities of a Case	Review paper	Justitie
33. Eckhouse, L.	2018	Layers of bias: a unified approach for understanding problems with risk assessment	Literatuuronderzoek	AI
34. Farzinder	2009	Machine translation of legal information and its evaluation	Experiment	Justitie
35. Fazelpour, S.	2021	Algorithmic bias: senses, sources, solutions	Literatuuronderzoek	AI
36. Fortes	2020	Paths to Digital Justice: Judicial Robots, Algorithmic Decision-Making, and Due Process	Case studie	Justitie
37. Franssen	2021	The use of AI tools in criminal courts: justice done and seen to be done?	Review paper	Justitie
38. Garingan, D.	2021	Artificial intelligence in legal practice: exploring theoretical frameworks for algorithmic literacy in the legal information provision	Literatuuronderzoek	Justitie
39. Giannini	2021	Lombroso 2.0: AI and predictions of dangerousness in criminal justice	Review paper	Justitie



40. Gillingham, P.	2019	Decision support systems, social justice and algorithmic accountability in social work: a new challenge	Literatuuronderzoek	Justitie
41. Green	2020	The false promise of risk assessment epistemic reforms and the limits of fairness	Review paper	Justitie
42. Green, B.	2019	Disparate interactions: an algorithm-in-the-loop analysis of fairness in risk assessments	Experiment	AI
43. Haarsma, G.	2020	Assessing risk among correctional community probation populations: predicting reoffense with mobile neurocognitive assessment software	Experiment	AI
44. Hälterlein, J.	2021	Epistemologies of predictive policing: mathematical social science, social physics and machine learning	Literatuuronderzoek	AI
45. Hamiltion	2018	The sexist algorithm	Experiment	Justitie
46. Hamilton, M.	2021	Evaluating algorithmic risk assessment	Literatuuronderzoek	AI
47. Hamilton, M.	2020	Judicial gatekeeping on scientific validity with risk assessment tools	Literatuuronderzoek	AI
48. Haque	2020	Understanding law enforcement and common peoples' perspective and designing explainable crime mapping algorithms	Experiment	Justitie
49. Harkens	2017	The ghost in the legal machine: algorithmic governmentality, economy and the practice of law	Review paper	Justitie
50. Hartmann	2021	Uncertainty, risk and the use of algorithms in policy decisions: a case study on criminal justice in the USA	Case studie	Justitie
51. Hayward, K. J.	2021	Artificial intelligence and crime: a primer for criminologists	Literatuuronderzoek	Justitie
52. He, P.	2011	A design of criminal investigation expert system based on CILS	Literatuuronderzoek	Justitie
53. Hilt	2017	What does the future hold for the law librarian in the advent of artifical intelligence	Literatuuronderzoek	Justitie
54. Hogan	2021	On the Ethics and Practicalities of Artificial Intelligence, Assessment, and Race	Review paper	GGZ
55. Hokkanen, J.	2002	Knowledge tools for legal knowledge tool makers	Literatuuronderzoek	AI
56. Huq	2019	Racial equity in algorithmic criminal justice	Review paper	Justitie
57. Karsai, K.	2021	Algorithmic decisions within the criminal justice ecosystem and their problem matrix	Literatuuronderzoek	Justitie
58. König	2021	When politization stops algorithms in criminal justice	Case studie	Justitie
59. König, P. D.	2021	Evaluating the evidence in algorithmic evidence-based decision-making: the case of US pretrial risk assessment tools	Experiment	AI
60. Kouli, R.	2020	Human control over automation: EU policy and AI ethics	Review paper	AI

61. Land, M. K.	2020	Human rights and technology: new challenges for justice and accountability	Literatuuronderzoek	
62. Lee, Y.	2019	A theory-driven algorithm for real-time crime hot spot forecasting	Literatuuronderzoek	Justitie
63. Leith	1988	The application of AI to law	Review paper	Justitie
64. Levy, K.	2021	Algorithms and decision-making in the public sector	Literatuuronderzoek	AI
65. Lin	2020	The limits of human predictions of recidivism	Experiment	Justitie
66. Liu	2019	Semantics and Structure Based Recommendation of Similar Legal Cases	Experiment	Justitie
67. Ludwig, J.	2021	Fragile algorithms and fallible decision-makers: lessons from the justice system	Literatuuronderzoek	Justitie
68. Mao, X.	2021	A prison term prediction model based on fact descriptions by capturing long historical information	Experiment	Justitie
69. Marjanovic, O.	2021	Algorithmic pollution: making the invisible visible	Literatuuronderzoek	Justitie
70. McKay	2019	Predicting risk in criminal procedure: actuarial tools, algorithms, AI and judicial decision-making	Review paper	Justitie
71. Mok	2021	Sentence Classification for Contract Law Cases: A Natural Language Processing Approach	Review paper en experiment	Justitie
72. Morison	2019	Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making	Review paper	Justitie
73. Moses, L. B.	2017	Artificial intelligence in the courts, legal academia and legal practice	Literatuuronderzoek	AI
74. Mothukuri, R.	2019	Data mining on prediction of crime and legal judgements: a state of an art	Literatuuronderzoek	Justitie
75. Müller, P.	2019	Algorithmic risk assessment als medium des rechts	Literatuuronderzoek	Justitie
76. O'Donnell, R. M.	2019	Challenging racist predictive policing algorithms under the equal protection clause	Literatuuronderzoek	Justitie
77. Oswald	2020	Technologies in the twilight zone: early lie detectors, machine learning and reformist legal realism	Review paper	Justitie
78. Park, M.	2021	AI model for predicting legal judgements to improve accuracy and explainability of online privacy invasion cases	Experiment	Justitie
79. Park, S-H.	2021	A survey of research on data analytics-based legal tech	Systematic review	Justitie
80. Peeters	2018	Machine justice: governing security through the bureaucracy of algorithms	Review paper	Justitie
81. Pelham III, W. E.	2020	Can machine learning improve screening for targeted delinquency prevention programs?	Experiment	Justitie

82. Pena, A. M. N.	2021	Artificial intelligence and precautionary guardianship: special reference to provisional prison	Literatuuronderzoek	Justitie
83. Piano, S. L.	2020	Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward	Literatuuronderzoek	Justitie
84. Prakken, H.	unknown	Legal reasoning: computational models	Literatuuronderzoek	Justitie
85. Quattrocolo	2019	Technical solutions for legal challenges: equality of arms in criminal proceedings	Case studies	Justitie
86. Resnick	2019	Passive Monitoring of Mental Health Status in the Criminal Forensic Population	Review paper	GGZ
87. Ruhl	2018	Topic modeling the president: conventional and computational methods	Review paper	Justitie
88. Scurich	2020	Public's Views of Risk Assessment Algorithms and Pretrial Decision Making	Experiment	Justitie
89. Shapiro	2019	Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing	Etnografie	Justitie
90. SIGAI group	unknown	Dutch artificial intelligence manifesto	Review paper	AI
91. Sil, R.	2019	Artificial intelligence and machine learning based legal application: the state-of-the-art and future research trends	Literatuuronderzoek	Justitie
92. Simmler	2021	Smart criminal justice: phenomena and normative requirements	Systematic review	Justitie
93. Skeem	2020	Using algorithms to address trade-offs inherent in predicting recidivism	Experiment	Justitie
94. Soukupová, J.	2021	AI-based legal technology: a critical assessment of the current use of artificial intelligence in legal practice	Literatuuronderzoek	Justitie
95. Stockhdale, M.	2019	Legal advice privilege and artificial legal intelligence: can robots give privileged legal advice?	Literatuuronderzoek	Justitie
96. Sukanya, G.	2021	A meta analysis of attention models on legal judgment prediction system	Systematic review	Justitie
97. Swofford	2022	Probabilistic reporting and algorithms in forensic science: Stakeholder perspectives within the American criminal justice system	Interviews	Justitie
98. Swofford, H.	2021	Implementation of algorithms in pattern & impression evidence: a responsible and practical roadmap	Literatuuronderzoek	AI
99. Ting	2018	Predicting recidivism among youth offenders: Augmenting professional judgement with machine learning algorithms	Experiment	Reclassering
100. Totorá	2020	Neuroprediction and A.I. in Forensic Psychiatry and Criminal Justice: A Neurolaw Perspective	Review paper	GGZ

101.	Vaccaro	2019	The effects of mixing machine learning and human judgement collaboration between human and machines does not necessarily lead to better outcomes	Experiment	Justitie
102.	Vaithianathan	2021	Using a machine learning tool to support high-stakes decision in child protection	Experiment	Justitie
103.	Van den Berg	2021	Validatiestudie risico en beschermende factoren van de RISC	Review paper	Reclassering
104.	Vincent	2020	Racist Algorithms or Systemic Problems?	Review paper	Justitie
105.	Watts	2020	Predicting offenses among individuals with psychiatric disorders - A machine learning approach	Experiment	GGZ
106.	Werth	2018	Risk and punishment: The recent history and uncertain future of actuarial, algorithmic, and "evidence-based" penal techniques	Review paper	Justitie
107.	Wilkins, D.	1997	The effectiveness of machine learning techniques for predicting time to case disposition	Experiment	Justitie
108.	Wright, S. A.	2020	AI in the law: towards assessing ethical risks	Literatuuronderzoek	Justitie
109.	Xivuri, K.	2021	A systematic review of fairness in artificial intelligence algorithms	Literatuuronderzoek	AI
110.	Zavrsnik, A.	2019	Algorithmic justice: algorithms and big data in criminal justice settings	Literatuuronderzoek	Justitie
111.	Zlibaite	2016	Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models	Case studies	Justitie

**Tabel 4.***Overzicht definities van AI*

<b>Auteurs</b>	<b>Publicatiejaar</b>	<b>Definitie van AI</b>
Xivuri & Twinomurinzi	2021	"AI enables systems to perform tasks that would normally be performed by humans in three different categories: narrow AI, general AI, and super AI. Narrow AI performs operational tasks through the use of machine learning tools such as recognising individual faces, driving a car, or speech recognition. General AI is designed to be as intelligent as humans with the ability to perform any intelligent tasks, but remains computationally complex. General AI solves complex problems and independently controls itself. General AI has the ability to get knowledge, apply it, reason, and think. Super AI is the type that is more intelligent than humans and would do better than humans in almost everything including intelligence and social skills. Super AI has not been developed yet and its implementation although being utopic is feared it could have negative consequences such as human extinction"
Sil	2019	"AI is based on the designing of smart computer systems that collects information, solves and makes decision and also act in an intelligence manner that is relatable to the intelligence of the human mind. AI is the study of system that extracts information from the surroundings and based on it performs the task. ML is an application of AI where the machines analyse data, learn patterns and assemble entire cognizance from the data by applying the rules (algorithms). It provides systems to optimize the ability to learn and improve the knowledge automatically."
Boahemaa	2018	"AI is science and a set of computational technologies inspired by people"
Hogan	2021	"AI performs functions heretofore limited to human intelligence"
SIGAI Group	2018	"AI science and engineering that studies and creates intelligent systems"
Bampasika	2020	"AI systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."
Alikhademi	2022	"Algorithms are used to automate tasks in a multitude of areas, such as health, finance, and law enforcement. Machine Learning (ML) is a class of algorithms that produces outcomes based on patterns found in data. It can be used for tasks ranging from cancer detection to predicting the likelihood that a parolee will re-offend. ML algorithms are gaining in popularity for tasks that can be time-consuming and cumbersome for humans to do. These algorithms produce results in less time and can overcome and mitigate human prejudices. However, ML methods can reflect and entrench the biases of humans."
O'Donnell	2019	"An algorithm is a specific sequence of logical operations that provides instructions for how computers should act on an input"

		data set. This specific sequence of logical operations is the source code. A machine learning algorithm learns from the training data that it is fed and finds correlational patterns within that data. The machine learning algorithm subsequently incorporates knowledge of these patterns into its code."
Hilt	2017	"By definition, cognitive computing contains "the fields of machine learning, reasoning and decision making technologies, language, speech and vision recognition and processing technologies, human interface technologies, and high performance computing"."
Soukupová	2021	Calo: "AI is a set of techniques aimed at approximating some aspect of human or animal cognition using machines", Scherer: "AI are machines that are capable of performing tasks that, if performed by a human, would be said to require intelligence". This is just a small demonstration of how diversely AI can be grasped. Thus, for the purpose of this article, artificial intelligence will be used as a broad umbrella term to cover a vast spectrum of technology often based on algorithms capable of achieving complex goals, irrespective of whether the technology is based on machine learning, natural language processing, deep learning or cognitive computing. This allows for an extensive amount of legaltech to be covered. AI-based legaltech then refers to tools designed to achieve certain goals to improve legal services while using artificial intelligence."
DoCarmo	2021	"Computational systems, Algorithmic governance on P4"
Dominic	2016	"From recent crime investigations, the crimes are being conducted by already registered criminals during their parole or coming out from jail after their sentence. In this system, a variety of photos related to each criminal is provided in the database. By using the Artificial Neural Network (ANN) algorithm the content of pictures that is input face, is automatically identified by the software. This approach of identifying images from a whole lot of images is used to understand the criminal from a wide collection of criminal lists. Using this software, it is easy to analyse the real details of one culprit through his face photo. Though this system provides a variety of appearance model photos about a criminal and categorizes them to a common criminal identification number, it is very easy to identify the criminal."
He	2011	"From the research achievements of AI and expert system (ES) in recent years (external sources), the original intention of researchers is that computers can substitute for the intelligence of human beings, thus acquire for the decision-making capacity of human experts and also overcome the limitation of experts in the field, so that to reach the level of true experts."
Beretta	2019	"Input, procedure and output"
Hayward	2021	"Machine learning (ML) involves the system gradually teaching itself the 'correct' (or 'useful') rules it needs to perform tasks effectively. Importantly, it does so on the basis of training data, rather than (as with 'expert systems') having these rules explicitly programmed. One specific type of ML, responsible for the current AI boom, is deep learning (DL). DL involves deep neural networks - an AI technique inspired by how neurons communicate with each other in biological brains. Artificial neural networks consist of layers of digital interconnected 'neurons', some of which receive an 'input' (e.g., information about a certain pixel in an image), while others provide an 'output' (e.g., a 'classification' of the image). Each neuron monitors others in the layer before it, and only if enough of those neurons send it a signal, will it then signal specific neurons in the subsequent layer. After each overall wrong/right response to the training data

		sample, the system alters the strength of the connections between the involved neurons. In this way, it 'learns' to encode the rules to do its task."
Das	2019	"Natural language processing techniques are mostly used for handling and processing these unstructured data by criminal investigators. Identifying named entities present in a text document helps in gaining knowledge about the persons involved in crime and discovering substantial relations among the identified named entities play an important part of taking proper actions in the criminal justice industry."
Bosse	2019	"Onderzoeken en ontwikkelen van intelligente systemen"
Deeks	2018	"Artificial intelligence is a notoriously capacious and slippery term. Generally, it refers to "a set of techniques aimed at approximating some aspect of human or animal cognition using machines.""
Moses	2017	"One kind of artificial intelligence technology used in law, the expert system, is designed to mimic the thought process of a human expert. <sup>13</sup> Expert systems rely on expertise drawn from the minds of human legal experts on a particular topic, either working with a 'knowledge engineer', relying on their own system design skills or deploying a software tool such as Neota Logic. They can be used in legal practice and legal academia. An expert system can ask the user questions to draw out a factual matrix or assess a student's knowledge. The expert system relies on the expertise with which it is programmed to provide advice or feedback. The conclusions of the system can be accompanied by reasons in the form of a series of logical statements and citations to relevant sources of authority."
Park	2021	"Recently, Industry 4.0 has received much attention from both researchers and practitioners. Industry 4.0 technologies, including artificial intelligence (AI), machine learning (ML), robotics, Internet of Things (IoT), wireless communication, big data, and clouds, are greatly affecting entire economics and society. Industry 4.0 technologies are recognized as powerful tools for innovating the productivity and competitiveness of a wide range of industries. The legal industry is no exception, and this paper focuses on the Industry 4.0 technologies applied to legal services. Traditionally, legal services are provided by human experts; however, modern technologies can be used to automate the service procedure of the legal domain."
Franssen	2021	"Software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal (...)."
Wilkins	1997	"The branch of artificial intelligence which investigates acquiring knowledge through observations called machine learning"
Mao	2021	"The legal judgement prediction (LJP) focus on applying artificial intelligence technology to provide decision support for legal workers. The prison term prediction (PTP) is an important task in LJP which aims to predict the term of penalty utilizing machine learning methods, thus supporting the judgement. The PTP is to support legal workers on predicting the term of penalty for a certain type of criminal case by analyzing the textual fact description."

**Tabel 5.**

Bestaande AI-applicaties in de internationale reclasseringspraktijk

<b>Auteur</b>	<b>Bicacki et al.</b>	<b>Berk</b>	<b>Bagaric et al.</b>	<b>Van den Berg</b>	<b>Ting et al.</b>
<i>Publicatiejaar</i>	2020	2017	2019		
<i>AI-applicatie</i>	Spatio-temporele patronen middels enkelbanden	Beslissingsondersteunend AI-systeem voor voorwaardelijk vrijkomen	Algoritmes risicotaxatie-instrumenten	in RISC, risicotaxatie-instrument binnen 3RO	Beslissingsondersteunend AI-systeem voor voorwaardelijk vrijkomen
<i>Reclasseringsproduct</i>	Toezichthouden	Toezichthouden	Advies en toezichthouden	Advies en toezichthouden	Toezichthouden
<i>Doel van de AI-applicatie</i>	Inzicht creëren in de beweegpatronen van de enkelbanddragers: waar en hoelang spenderen zij hun tijd?	Beslissen wie voorwaardelijk vrijkomt met toezicht en wie (weer) naar de gevangenis moeten	Het berekenen van de recidive-kans om een oordeel te kunnen geven over het voorwaardelijk vrijkomen of het (terug)plaatsen in de gevangenis	Het berekenen van de recidive-kans om een oordeel te kunnen geven over het voorwaardelijk vrijkomen of het (terug)plaatsen in de gevangenis	Beslissen wie voorwaardelijk vrijkomt met toezicht en wie (weer) naar de gevangenis moeten
<i>Belang voor de praktijk</i>	Mogelijkheid om per enkelbanddrager te evalueren met de reclasseringswerker welk gebied voor de enkelbanddrager het meest aantrekkelijk is. Tevens relevant voor beleidsmakers in het domein van toezichthouden van de reclassering mits de data uitgebreider wordt verzameld	Er werd meer onderscheid gemaakt door de beslissingscommissie met behulp van het AI-systeem tussen veroordeelden voor een geweldloos of gewelddadig delict. Ook waren de gemaakte beslissingen ondersteunt door AI accurater en correcter voor degene die geweldloze delicten pleegden	Een risicotaxatie is een accurater middel om de recidive-kans te berekenen vergeleken met een ongestructureerd oordeel van deze recidive-kans. De risicotaxatie-instrumenten kunnen ook ingezet worden om een individueel traject op te zetten binnen de reclassering	RISC wordt momenteel al toegepast binnen 3RO, en is bedoeld als ondersteuning voor de reclasseringswerker.	Door recidive te voorspellen, is het mogelijk om interventies op maat te maken
<i>Aandachtspunten</i>	Er wordt niet nader ingegaan op de voor- en nadelen van het gebruik van de data verkregen vanuit de enkelbanden op moreel, juridisch en ethisch vlak	Er wordt niet nader ingegaan op de voor- en nadelen van het gebruik van dit beslissingsondersteunend AI-systeem op moreel, juridisch en ethisch vlak	Algoritmen zijn ontworpen om op informatie in te spelen die zij toebedeeld krijgen en hebben geen weet over sociale acceptatie. Hierdoor kunnen algoritmen discriminerend recidive-kansen	Er moet gestreefd worden naar een zo hoog mogelijke overeenstemming tussen beoordelaars op alle niveaus van de RISC. Verder vervolg onderzoek is nodig naar de eindoordelen op basis van de RISC.	



toekennen aan  
cliënten. Het  
instrument dient ter  
ondersteuning van  
menselijke  
beslissingen (ten  
opzichte van het  
vervangen).

**Tabel 6.**

Bestaande AI-applicaties in de internationale reclasseringspraktijk

Predictive policing	Justitiële/rechterlijke beslissingen	Info	GGZ	Overig
1, 4, 14, 16, 37, 44, 48, 51, 56, 62, 76, 80, 89, 91, 92, 93,	2, 4, 12, 14, 15, 17, 20, 24, 26, 27, 28, 30, 32, 35, 37, 39, 40, 41, 43, 45, 46, 47, 50, 56, 57, 58, 59, 65, 66, 67, 68, 70, 72, 73, 74, 78, 80, 81, 82, 83, 84, 85, 88, 91, 94, 97, 99, 101, 102, 106, 107, 110,  [specifiek compass genoemd: 4, 12, 14, 15, 17, 30, 32, 35, 37, 45, 50, 65, 72, 82, 83, 94, 101, 110]	21, 22, 28, 31, 32, 34, 37, 38, 53, 55, 66, 71, 74, 79, 91, 96	54, 86, 105	2, 29, 52, 61, 67, 77, 94, 95, 98, 100, 102,

*Cijfers corresponderen met de lijst papers beschikbaar in tabel 3*

**Tabel 7.**

Genoemde voor- en nadelen aan het inzetten van AI binnen het justitiële veld

<b>Nadelen AI</b>			<b>Voordelen AI</b>			
<i>Bias</i>	<i>Black box</i>	<i>Acceptatie</i>	<i>Snel accurate beslissingen kunnen nemen</i>	<i>Onbevooroordeeld van beslissingen</i>		
1, 3, 4, 6, 7, 10, 16, 18, 23, 25, 27, 30, 33, 35, 36, 37, 39, 40, 41, 42, 45, 46, 47, 51, 54, 56, 57, 58, 60, 63, 64, 69, 70, 73, 76, 80, 82, 86, 88, 89, 92, 93, 94, 100, 101, 103, 104, 105, 106, ,109, 110, 111	1, 3, 4, 6, 17, 18, 23, 24, 25, 36, 37, 39, 40, 44, 57, 59, 60, 61, 64, 69, 76, 83, 92, 94, 102, 110,	3, 5, 4, 6, 14, 18, 20, 25, 70, 72, 84, 97, 110	2, 4, 8, 12, 19, 20, 23, 27, 33, 36, 37, 41, 47, 49, 50, 51, 52, 53, 54, 65, 67, 70, 80, 92, 94, 109	4, 16, 19, 27, 36, 37, 41, 47, 49, 50, 51, 54, 57, 65, 67, 70, 72, 73, 89, 109,		

Cijfers corresponderen met de lijst papers beschikbaar in tabel 3.

## Appendix B: Protocol interview AI-experts

*De verschillende manieren van gebruik van AI in het algemeen en de meeste passende definitie (onderzoeksvraag 1)*

1. Waar denkt u aan bij de term Artificial Intelligence?
  - a. Wat is volgens u een passende definitie voor Artificial Intelligence?

*De huidige toepassing van AI in de justitiële context en 3RO (onderzoeksvraag 2 en 3)\*

2. Bent u op de hoogte van het gebruik van AI binnen de reclassering (nationaal/internationaal)?
  - a. Zo ja, welke voorbeelden kent u?
  - b. In hoeverre is dit gebruik relevant?
  - c. In hoeverre is dit gebruik van toepassing voor de reclassering?
3. Bent u op de hoogte over hoe AI op dit moment binnen (gelijke) juridische praktijken wordt gebruikt (nationaal/internationaal)?
  - a. Zo ja, welke voorbeelden kent u?
  - b. In hoeverre is dit gebruik relevant?
  - c. In hoeverre is dit gebruik van toepassing voor de (gelijke) juridische praktijken?

Mogelijke ondersteuningsvragen op basis van de 5 3RO-taken:

4. Op welke manier(en) wordt Artificial Intelligence gebruikt in diagnose en advies aan het OM en de rechterlijke macht?
  - a. In hoeverre is de huidige toepassing van AI passend voor deze 3RO-taak?
5. Op welke manier(en) wordt Artificial Intelligence gebruikt in het toezien op de uitvoering van werkstraffen?
  - a. In hoeverre is de huidige toepassing van AI passend voor deze 3RO-taak?
6. Op welke manier(en) wordt Artificial Intelligence gebruikt in het toezicht houden op verdachten en/of daders?
  - a. In hoeverre is de huidige toepassing van AI passend voor deze 3RO-taak?
7. Op welke manier(en) wordt Artificial Intelligence gebruikt in het indiceren voor forensische zorg?
  - a. In hoeverre is de huidige toepassing van AI passend voor deze 3RO-taak?
8. Op welke manier(en) wordt Artificial Intelligence gebruikt in de totstandkoming van ambulante zorg of beschermd wonen voor forensische zorg?
  - a. In hoeverre is de huidige toepassing van AI passend voor deze 3RO-taak?

*Mogelijke toepassingsmogelijkheden van AI binnen de reclasseringspraktijk specifiek, en de voor- en nadelen hiervan (onderzoeksvraag 4 en 5)*

De volgende taken behoren tot de drie reclasseringsorganisaties (3RO): (a) Advies, (b) Toezicht, (c) Werkstraffen, (d) Justitiële interventies.

9. Hoe ziet u de toepassingsmogelijkheden van Artificial Intelligence binnen de reclasseringstaak 'Advies' voor u?
  - a. Wat kunnen voordelen zijn van deze toepassingsmogelijkheid/-heden?
  - b. Wat kunnen nadelen zijn van deze toepassingsmogelijkheid/-heden?
10. Hoe ziet u de toepassingsmogelijkheden van Artificial Intelligence binnen de reclasseringstaak 'Toezicht voor u'
  - a. Wat kunnen voordelen zijn van deze toepassingsmogelijkheid/-heden?
  - b. Wat kunnen nadelen zijn van deze toepassingsmogelijkheid/-heden?

11. Hoe ziet u de toepassingsmogelijkheden van Artificial Intelligence binnen de reclasseringstaak 'Werkstraffen' voor u?
  - a. Wat kunnen voordelen zijn van deze toepassingsmogelijkheid/-heden?
  - b. Wat kunnen nadelen zijn van deze toepassingsmogelijkheid/-heden?
12. Hoe ziet u de toepassingsmogelijkheden van Artificial Intelligence binnen de reclasseringstaak 'Justitiële interventies voor u'?
  - a. Wat kunnen voordelen zijn van deze toepassingsmogelijkheid/-heden?
  - b. Wat kunnen nadelen zijn van deze toepassingsmogelijkheid/-heden?

*Morele, ethische en juridische aspecten die de rol van AI binnen de reclasseringspraktijk kunnen beïnvloeden (onderzoeksvraag 6)*

Met de input van toepassingsmogelijkheden voortkomend uit de voorgaande vragen, wordt er per idee uitgevraagd naar ethische/morele en juridische aspecten omtrent mogelijke invloeden in de rol van Artificial Intelligence binnen de reclasseringspraktijk.

13. Welke ethische/morele aspecten kunnen de rol van Artificial Intelligence van ... *(vul in: het idee over de toepassingsmogelijkheid)* binnen de reclasseringspraktijk beïnvloeden?
14. Welke juridische aspecten kunnen de rol van Artificial Intelligence van ... *(vul in: het idee over de toepassingsmogelijkheid)* binnen de reclasseringspraktijk beïnvloeden?

*Vervolgstappen: wat is er nodig om AI succesvol in te zetten in de reclasseringspraktijk? (onderzoeksvraag 7 en 8)*

Met de input van de nadelen omtrent de toepassingsmogelijkheden voortkomend uit de voorgaande vragen, wordt er per nadeel uitgevraagd wat er voor nodig is om deze barrières te overkomen en het werkbaar te maken voor de reclasseringspraktijk.

15. Wat is er nodig om ... *(vul in: het nadeel van de toepassingsmogelijkheid)* op te lossen?

Met de input van de meest wenselijke toepassingsmogelijkheid wordt er uit gevraagd wat er nodig is om deze AI-toepassing succesvol in te zetten in de reclasseringspraktijk.

16. Wat is er nodig om ... *(vul in: meest wenselijke toepassingsmogelijkheid)* succesvol in te zetten in de reclasseringspraktijk?

## Appendix C: Protocol interview reclasseringswerkers

Welke kennis is er momenteel op het gebied van AI? (onderzoeksvraag 2)

1. Heeft u voor dit onderzoek al eens gehoord van het begrip AI?
  - Zo ja, wat kunt u hierover vertellen (algemeen en reclassering)?
  - Zo nee, wat is het eerste waar u aan denkt bij het horen van het begrip AI?

Programma's die momenteel in gebruik zijn door de reclasseringswerkers die als AI te classificeren zijn (onderzoeksvraag 3)

Korte uitleg van AI: AI wordt beschouwd als een set van technologieën die zorgen voor het verzamelen, analyseren en presenteren van een grote hoeveelheid data op een snelle manier. Met AI wordt er gepoogd om menselijke taken uit te voeren en te automatiseren. Het kan ons leven makkelijker maken door bijvoorbeeld snel onderbouwde beslissingen te nemen.

1. Welke programma's die momenteel in gebruik zijn door u als reclasseringswerker zijn gemaakt op basis van AI die bij u bekend zijn?

Huidige programma's die momenteel in gebruik zijn door reclasseringswerkers die als AI te classificeren zijn:

- Advies:
  - a. RISC (OxRec en Verdiepende instrumenten: Zeden, Relatieel geweld, Gewelddadig extremisme, Licht verstandelijke beperking)
  - b. ZSM (Zorgvuldig, Snel en op Maat)
- Toezicht: Elektronische monitoring (EM)
  - a. RFID-enkelband
  - b. GPS-enkelband
  - c. Alcoholmeter
- Justitiële interventies: Stap voor Stap app

Bij ieder AI-geclassificeerd programma worden de volgende vragen besproken:

2. Hoe ervaart u het huidige gebruik van deze programma's in uw dagelijkse werk binnen de reclasseringspraktijk?
  - a. Wat zijn voordelen van het AI-geclassificeerde programma?
  - b. Wat zijn nadelen van het AI-geclassificeerde programma?

Visie op het gebruik van AI binnen de reclasseringspraktijk (onderzoeksvraag 5 en 6)

1. In wat voor taken behorende bij de functie als reclasseringswerker ziet u een mogelijkheid om een AI-toepassing toe te voegen? Bijvoorbeeld: denk aan uw belangrijkste taken in uw werk. Waar zou AI daar bij kunnen helpen?

De volgende mogelijke toepassingen van AI binnen de reclasseringspraktijk worden voorgelegd aan de reclasseringsmedewerker:

1. AI om risico's in te schatten en/of voorspellen
2. AI om informatie te ordenen
3. Interactieve AI (chatbots etc.)

Bij iedere toepassingsmogelijkheid worden de volgende vragen besproken:

1. Wat zijn voordelen van mogelijk gebruik van deze AI-toepassing voor u als reclasseringsmedewerker?
2. Wat zijn nadelen van mogelijk gebruik van deze AI-toepassing voor u als reclasseringsmedewerker?
3. Hoe zou de balans tussen beslissingen vanuit AI versus die door de reclasseringswerkers eruit moeten zien? Hoeveel verantwoordelijkheid *kan* en *mag* er bij AI komen te liggen?

Additionele vraag: Zijn er (nu) nog andere mogelijke AI-toepassingen waar u aan denkt bij de 4 taken als reclasseringswerker?

(Optioneel)

<u>Advies</u>
<ul style="list-style-type: none"><li>&gt; Machine learning om de juiste informatie te halen uit multimodale bronnen (foto's, tekst en video)</li><li>&gt; Algoritmen om te redeneren op basis van informatie in bijvoorbeeld (juridische) documenten en opgestelde misdaadscenario's</li><li>&gt; Simulaties van complexe criminele systemen en robotica</li><li>&gt; Algoritmen ter ondersteuning: voorkeurinterventie en alternatieve interventie aanbieden op basis van real-time data dat inzicht geeft wat er speelt</li><li>&gt; Gesproken woorden op geluids- of beeld dragers worden herkend en omgezet naar geschreven tekst en vervolgens geanalyseerd</li><li>&gt; AI-systeem waarin alle relevante data wordt gecombineerd en de analyse leidt tot een aanbeveling aan de rechter bijvoorbeeld</li></ul>
<u>Toezicht</u>
<ul style="list-style-type: none"><li>&gt; Toekomstig (crimineel) gedrag voorspellen op basis van data verkregen door elektronisch monitoring: wat is de kans dat iemand een criminele activiteit passeert?</li><li>&gt; Begeleiden en controleren van toezicht: elektronisch monitoring data gebruiken om het EM-gedrag te bespreken (en dus uiteindelijk te interveniëren)</li></ul>
<u>Werkstraffen</u>
<ul style="list-style-type: none"><li>&gt; Algoritme die redeneert of er een match is tussen de cliënt en het werkstrafproject</li></ul>
<u>Justitiële interventies</u>
<ul style="list-style-type: none"><li>&gt; Chatbots en robots voor gedragsverandering binnen 3RO (bijvoorbeeld sociale robot Maatje)</li><li>&gt; Virtuele coaches (computer vs. mens gestuurd)</li><li>&gt; Algoritmen ter ondersteuning: voorkeurinterventie en alternatieve interventie aanbieden op basis van real-time data dat inzicht geeft wat er speelt</li></ul>

*Visie op vervolgstappen binnen 3RO om AI toe te kunnen passen binnen de reclasseringspraktijk (onderzoeksvraag 7 en 8)*

1. Hoe denkt u over de groeiende rol in het gebruik van AI-geclassificeerde programma's in de functie als reclasseringswerker binnen de reclasseringspraktijk?
  - a. Ziet u (groei)mogelijkheden in de groeiende rol in het gebruik van AI-geclassificeerde programma's binnen de reclasseringspraktijk?
  - b. Ziet u valkuilen in de groeiende rol in het gebruik van AI-geclassificeerde programma's binnen de reclasseringspraktijk?
2. Wanneer kan AI succesvol worden toegepast binnen de reclasseringspraktijk volgens u?
  - a. Wat zijn hierbij aandachtspunten volgens u?