

Wat valt onder de AI-verordening?

Een analyse naar aanleiding van het inspectierapport Risicovol Algoritmegebruik over risicotaxatie bij de reclassering

Johannes Bijlsma, Floris Bex, Gerben Meynen & Maaïke Kempes¹

Met de komst van de Europese AI-verordening (AI Act) is het belangrijk vast te stellen welke systemen onder deze regelgeving vallen. In het bijzonder roept de toepassing van algoritmes binnen domeinen zoals de strafrechtspleging discussie op wanneer gebruik wordt gemaakt van klassieke statistische modellen. In dit artikel verkennen de auteurs, aan de hand van het voorbeeld van risicotaxatie-instrumenten zoals *OxRec*, waar de grens ligt tussen traditionele statistiek en artificiële intelligentie, en welke gevolgen deze afbakening heeft voor de reikwijdte van de AI-verordening.

1. Inleiding

Op 12 februari 2026 publiceerde de Inspectie Justitie en Veiligheid het rapport *Risicovol algoritmegebruik: Onderzoek naar gebruik van algoritmes bij de reclassering*.² In dit rapport onderzocht de inspectie of algoritmes bij de reclassering op een verantwoorde manier worden gebruikt.

De reclassering gebruikt software die algoritmes bevat waarmee aan de hand van risicofactoren (bijvoorbeeld leeftijd, geslacht, justitiële voorgeschiedenis) een inschatting wordt gemaakt van de kans dat een verdachte of veroordeelde zal recidiveren. Deze inschatting wordt onder meer gebruikt om de rechter te adviseren. De inspectie oordeelde dat één instrument dat de reclassering gebruikt om het risico op recidive van verdachte personen in te schatten, de *Oxford Risk of Recidivism Tool (OxRec)*, voldoet aan de definitie van AI-systeem in de AI-verordening (ook wel: AI Act).³ Over drie andere risicotaxatie-instrumenten oordeelde de inspectie dat het geen AI-systemen zijn.⁴ Het betreft de instrumenten *STATIC-99R*, *STABLE-2007* en *ACUTE-2007*, die gebruikt worden om het risico op recidive van mannelijke zedendelinquenten in te schatten.

AI-systemen die gebruikt worden om te beoordelen hoe groot het risico is dat iemand (opnieuw) een strafbaar feit zal plegen zijn volgens de AI-verordening 'hoogrisico-systemen'.⁵ Daarvoor gelden strenge eisen en in bepaalde gevallen – bijvoorbeeld als geen sprake is van een mense-

lijke eindbeoordeling – zijn deze zelfs verboden.⁶ Volgens de inspectie is *OxRec* een dergelijk hoogrisico AI-systeem. Het aanmerken van een algoritme als AI-systeem in de zin van de AI-verordening heeft belangrijke gevolgen. In dat geval moet aan eisen worden voldaan op het gebied van onder meer transparantie, menselijk toezicht, robuustheid en grondrechtenbeoordelingen.

De Autoriteit Persoonsgegevens wordt in het recent in internetconsultatie gebrachte wetsvoorstel tot uitvoering van de Verordening artificiële intelligentie als markttoezichtautoriteit aangemerkt voor hoogrisicotoepassingen.⁷ Desalniettemin heeft de Inspectie Justitie en Veiligheid als onafhankelijke toezichthouder een gezaghebbende positie, waardoor haar oordelen richtinggevend zijn voor beleid en praktijk in de rechtshandhaving, ook buiten de specifieke casus die zij onderzoekt. Dat geldt ook voor haar oordeel dat *OxRec* een AI-systeem is in de zin van de AI-verordening,⁸ en de drie andere instrumenten niet. De argumenten die de inspectie en de reclassering daarover hebben uitgewisseld – en die in het rapport en de bijbehorende hoor-en-wederhoorprocedure zijn gedocumenteerd⁹ – verdienen daarom nadere beschouwing. Deze discussie raakt namelijk aan de bredere, onderliggende vraag: wat is de reikwijdte van de AI-verordening?

OxRec maakt gebruik van klassieke statistische technieken. Dergelijke technieken worden al decennialang toegepast binnen de strafrechtsketen, niet alleen voor het

inschatten van recidiverisico's, maar bijvoorbeeld ook als basis voor psychologische tests die in het kader van pro Justitia-rapportages worden afgenomen. Algoritmes spelen in deze psychologische tests vooral een rol bij het omzetten van antwoorden naar scores en het vergelijken van deze scores met normgroepen. Op basis van vooraf vastgestelde statistische modellen worden resultaten gestandaardiseerd en geïnterpreteerd, bijvoorbeeld om bepaalde persoonlijkheidskenmerken of cognitieve eigenschappen te kwantificeren. In pro Justitia-rapportages ondersteunt dit de beoordeling van de kans op recidive. De uiteindelijke duiding is aan de deskundige.

Vergelijkbare, op statistiek gebaseerde instrumenten worden echter veel breder ingezet voor, bijvoorbeeld, het voorspellen van studie-uitval in het onderwijs, het rangschikken van sollicitanten op de arbeidsmarkt, het beoordelen van kredietwaardigheid bij de toegang tot financiële diensten, en het berekenen van risicoscores voor hart- en vaatziekten in de geneeskunde. Hoewel niet al deze domeinen onder de hoogrisicosystemen van de AI-verordening vallen, zou het aanmerken van klassieke statistische methoden als 'AI' toch een groot deel van de alledaagse praktijk in deze domeinen binnen het bereik van de verordening brengen. Dit roept belangrijke afbakeningsvragen op. De verordening was immers bedoeld om kunstmatige intelligentie te reguleren, niet de statistiek. Tegelijkertijd is *OxRec* software die een risicoscore genereert die van invloed kan zijn op de vrijheid van een persoon en is de AI-verordening mede ontworpen om voorspellende systemen in een dergelijke context te reguleren.¹⁰

In dit artikel onderzoeken wij aan de hand van de door de reclassering gebruikte risicotaxatie-instrumenten, in het bijzonder *OxRec*, welke argumenten er zijn om klassieke statistiek gebaseerde algoritmes als AI-systeem in de zin van de AI-verordening aan te merken. Daartoe bespreken we ten eerste beknopt de definitie van AI-systeem in de zin van de AI-verordening. Ten tweede richten we ons op de vraag wanneer risicotaxatie-instrumenten over aanpassingsvermogen en inferentievermogen beschikken – twee centrale elementen in de definitie van

Welke argumenten er zijn om op klassieke statistiek gebaseerde algoritmes als AI-systeem in de zin van de AI-verordening aan te merken

AI-systeem in de verordening. We sluiten af met enkele belangrijke discussiepunten. Blijken zal dat er nog veel onduidelijkheden zijn, waarmee dit artikel ook gezien kan worden als een eerste verkenning.

2. De definitie van AI-systeem in de AI-verordening

De AI-verordening bevat regels voor de ontwikkeling en het gebruik van AI-systemen met onder andere het oog op de bescherming van grondrechten en de rechtsstaat en ter bevordering van innovatie.¹¹ De verordening stelt dwingende eisen aan risicobeheersing, databeheer, transparantie, menselijk toezicht en beveiliging voor de ontwikkeling en het gebruik van AI-systemen met een hoog risico. Ook moet het systeem zelf voldoende transparant, nauwkeurig en robuust zijn.¹²

In de verordening wordt vooropgesteld dat het begrip 'AI-systeem' duidelijk moet worden gedefinieerd met het oog op rechtszekerheid en tegelijkertijd de nodige flexibiliteit moet bieden om in te kunnen spelen op nieuwe technologische ontwikkelingen.¹³ 'AI-systeem' wordt in artikel 3 lid 1 AI-verordening als volgt gedefinieerd: 'een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen'.

Auteurs

1. Prof. mr. dr. J. Bijlsma is universitair hoofddocent strafrecht, Utrecht Centre for Accountability and Liability Law (UCALL) en Willem Pompe Instituut voor Strafrechtswetenschappen (WPPI), Universiteit Utrecht (UU), en bijzonder hoogleraar strafrechtswetenschappen, vakgroep Transboundary Legal Studies, Rijksuniversiteit Groningen. Prof. dr. F.J. Bex is hoogleraar innovatie in artificiële intelligentie en recht, departement Informatica/departement Rechtsgeleerdheid, UU. Prof. dr. G. Meynen is hoogleraar forensische psychiatrie, UCALL en WPPI, UU en hoogleraar ethiek, in het bijzonder bio-ethiek, afdeling Filosofie, faculteit Sociale en Geesteswetenschappen, Vrije Universiteit Amsterdam. Prof. dr. M. Kempes is hoofd wetenschap en opleidingen bij het Nederlands Instituut voor Forensische Psychiatrie

en Psychologie en bijzonder hoogleraar forensische neuropsychologie, Instituut voor Pedagogische Wetenschappen, faculteit Sociale Wetenschappen, UU.

Noten

2. Inspectie Justitie en Veiligheid, *Risicovol algoritmegebruik: Onderzoek naar gebruik van algoritmes bij de reclassering*, 2026.
3. Verordening (EU) 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie en tot wijziging van de Verordeningen (EG) 300/2008, (EU) 167/2013, (EU) 168/2013, (EU) 2018/858, (EU) 2018/1139 en (EU) 2019/2144, en de Richtlijnen 2014/90/EU, (EU) 2016/797 en (EU) 2020/1828 ('AI-verordening').
4. *Risicovol algoritmegebruik*, p. 27-28.

5. Bijlage III, onder 6. d) AI-verordening.

6. Art. 5 lid 1 onder d AI-verordening.

7. 'Uitvoeringswet AI-verordening', internetconsultatie.nl/uaiv/b1.

8. De voorzitter van de Autoriteit Persoonsgegevens noemde naar aanleiding van het rapport van de inspectie OxRec als een voorbeeld van een AI-systeem ('Autoriteit Persoonsgegevens wil snel AI-regels: "Grondrechten moeten beschermd"', nos.nl 5 maart 2026.)

9. 'Hoor-wederhoor reactie Reclassering', inspectie-jenv.nl.

10. Bijlage III, 6 onder d AI-verordening noemt expliciet: 'AI-systemen die bedoeld zijn om door of namens rechtshandavingsinstanties of door instellingen, organen of instanties van de Unie ter ondersteuning van rechtshandavingsinstanties te worden gebruikt om te beoordelen hoe groot het

risico is dat een natuurlijke persoon (opnieuw) een strafbaar feit zal plegen.'

11. Overweging 1 AI-verordening. Zie voor een beschrijving en analyse van (de totstandkoming van) de AI-verordening: N.A. Smuha & K. Yeung, 'The European Union's AI Act. Beyond motherhood and apple pie?', in: N.A. Smuha e.a. (red.), *The Cambridge handbook of the law, ethics and policy of artificial intelligence*, Cambridge: Cambridge University Press 2025, p. 228-258. De spanning tussen regulering en flexibiliteit komt duidelijk naar voren in de recente politieke ontwikkelingen rondom het vereenvoudigen van de eisen voor 'industriële AI' ('"Ondemocratische draai" en uitstelgedrag rond de AI Act laten Brusselse worsteling met deregulering en de VS zien', NRC 7 mei 2026.

12. Afdeling 2 AI-verordening.

13. Overweging 12 AI-verordening.



© Getty images

Deze complexe definitie wordt in overweging 12 van de verordening ingekleurd.¹⁴ Ook heeft de Europese Commissie conform artikel 96 AI-verordening 'richtsnoeren' (*guidelines*) gepubliceerd waarin getracht wordt het begrip verder te verduidelijken.¹⁵ Deze richtsnoeren zijn niet bindend en kunnen evolueren, maar kunnen wel indicatief zijn voor hoe de Commissie en toezichhoudende autoriteiten om zouden kunnen gaan met de vraag welke systemen als AI kunnen worden aangemerkt. Zowel de inspectie als de reclassering baseren zich op deze richtsnoeren in hun debat over de vraag of *OxRec* een AI-systeem in de zin van de verordening is. Aangezien het niet mogelijk is om binnen het bestek van dit artikel op alle aspecten van de definitie in te gaan, richten wij ons op de twee kernelementen: aanpassingsvermogen en inferentievermogen.

3. Zijn risicotaxatie-instrumenten AI-systemen?

3.1. Heeft *OxRec* aanpassingsvermogen?

Als eerste is het belangrijk om kort het aanpassingsvermogen van AI zoals genoemd in de definitie te bespreken, omdat dit een mogelijke bron van verwarring is. Zoals in overweging 12 van de verordening genoemd wordt, heeft aanpassingsvermogen betrekking op 'zelflerende capaciteiten, waardoor het systeem tijdens het gebruik kan veranderen'. Echter, systemen die zich niet aanpassen na ingebruikname kunnen ook AI zijn. Het aanpassingsvermogen dat in de definitie van AI-systeem is opgenomen is namelijk geen noodzakelijke voorwaarde om aan de definitie te voldoen. Dat komt tot uitdrukking in de clausule-

ring dat een AI-systeem aanpassingsvermogen *kan* vertonen na inzet.¹⁶

De meeste *machine learning*-systemen die op dit moment in gebruik zijn werken als volgt. Een algoritme wordt getraind op een dataset van gevallen waarvan de uitkomst bekend is, bijvoorbeeld een groep voormalige gedetineerden van wie bekend is of zij na vrijlating wel of niet hebben gerecidiveerd. Het algoritme leert welke combinatie van kenmerken, zoals leeftijd, geslacht en eerdere veroordelingen, samenhangt met recidive. Op basis daarvan legt het een relatie tussen de kenmerken (input) en de kans of het risico op recidive (output). Die relatie wordt uit de data geleerd waarna het algoritme na ingebruikname zonder verdere aanpassing wordt toegepast op nieuwe gevallen.¹⁷ Van continu leren *tijdens* het gebruik is bij de meeste *machine learning*-systemen dus geen sprake. Dat is voorbehouden aan een beperkte categorie systemen die na elke interactie automatisch worden bijgesteld, zoals de productaanbevelingen van onlinewinkels die direct veranderen op basis van wat een gebruiker bekijkt of koopt.

De reclassering en de inspectie zijn het eens over het feit dat *OxRec* zich tijdens de gebruiksfase niet meer aanpast. *OxRec* is dus niet continu zelflerend.¹⁸ Dat een risicotaxatie-instrument na ingebruikname niet verder leert, betekent evenwel niet dat het geen AI-systeem in de betekenis van de verordening kan zijn.

3.2. Heeft *OxRec* inferentievermogen?

Inferentievermogen – het in artikel 3 lid 1 AI-verordening omschreven vermogen om 'uit de ontvangen input [af te leiden] hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen' – is een essentieel

kenmerk van een AI-systeem. Op zichzelf is dit echter heel algemeen geformuleerd. Ook een rekenmachine genereert immers output op basis van input.¹⁹ Overweging 12 verduidelijkt daarom dat het bij inferentievermogen moet gaan om iets dat de elementaire gegevensverwerking zoals traditionele softwaresystemen die uitvoeren *overstijgt*, door leren, redeneren of modelleren mogelijk te maken. Inferentievermogen kan bovendien betrekking hebben op het vermogen van AI-systemen om modellen of algoritmen, of beide, af te leiden uit input of data.²⁰

OxRec is een actuair risicotaxatie-instrument dat wordt gebruikt om de kans in te schatten dat iemand opnieuw een strafbaar feit pleegt.²¹ Het model is ontwikkeld in het Verenigd Koninkrijk en is gebaseerd op statistische analyse van grote datasets met strafrechtelijke gegevens. Het gebruikt vooral statische, niet te veranderen, factoren zoals: leeftijd, geslacht, eerdere veroordelingen en type delict. Op basis hiervan berekent *OxRec* een risicoscore voor recidive. *OxRec* is bedoeld om de advisering van de reclassering aan de rechter te ondersteunen en wordt ook gebruikt in het kader van detentiefasering en reclasseringstoezicht, maar vervangt de professionele beoordeling niet.²² Het model waarop *OxRec* gebaseerd is heeft door middel van *Cox proportional hazards*-regressie de relatie bepaald tussen input (risicofactoren) en output (kans op recidive).²³ *Cox proportional hazards*-regressie is een statistische methode die in de jaren zeventig van de vorige eeuw door David Cox is ontwikkeld om overlevingskansen van patiënten te berekenen.²⁴

Zo bezien zou gezegd kunnen worden dat *OxRec* over inferentievermogen beschikt: net als de *machine learning*-systemen die in paragraaf 3.1 werden beschreven bepaalt – of ‘leert’ – *OxRec* zelf de relatie tussen input en output uit historische recidivedata. Met andere woorden, het leidt

een model of algoritme af uit data. Volgens de inspectie is daarom sprake van inferentievermogen:

‘Er is sprake van *supervised machine learning*, omdat het [*OxRec*]-algoritme bij de ontwikkeling getraind is op de gelabelde data van 37.100 Zweedse gedetineerden. De labels hierbij zijn of de gedetineerden wel of niet gerecidiveerd hebben en wanneer dit plaatsvond. Er is door middel van *Cox proportional hazards regression* bepaald wat de relatie is tussen de input variabelen en de kans op recidive. Hiermee vertoont het systeem het vermogen om te infereren.’²⁵

De redenering van de inspectie is op het eerste gezicht overtuigend: *OxRec* is immers getraind op gelabelde historische data. Zoals we eerder zagen werkt *supervised machine learning* op dezelfde wijze.²⁶ Betoogd kan daarom inderdaad worden dat *OxRec* een zekere mate van inferentievermogen heeft. Toch zijn er twee argumenten om een slag om de arm te houden.

Ten eerste zou *OxRec* gezien kunnen worden als een ‘eenvoudig voorspellingsstelsel’ dat volgens de huidige richtsnoeren buiten het toepassingsgebied van de AI-verordening valt.²⁷ Van dit soort systemen is sprake wanneer een *machine learning*-model in feite niet meer doet dan wat ook met een eenvoudige vuistregel bereikt had kunnen worden, zoals ‘neem altijd het gemiddelde’. De richtsnoeren geven als voorbeelden het voorspellen van de temperatuur van morgen op basis van het gemiddelde van de vorige week, of het schatten van de dagelijkse vraag naar een product in een winkel op basis van eerdere gemiddelden. *OxRec* lijkt echter niet onder deze uitzondering te vallen: het modelleert via *Cox proportional hazards*-regressie de samenhang tussen meerdere

OxRec is niet continu zelflerend. Dat een risicotaxatie-instrument na ingebruikname niet verder leert, betekent evenwel niet dat het geen AI-systeem in de betekenis van de verordening kan zijn

14. Sinds het begin van het wetgevingstraject is de definitie een aantal keer aangepast, mede om haar meer in lijn te brengen met de definitie die de OESO hanteert. Zie over de ontwikkeling van het begrip: *Commission guidelines on the application of the definition of an AI system and the prohibited AI practices established in the AI Act. Response of the European Law Institute*, Wenen: ELI 2024, p. 9-10; Smuha & Yeung 2025, p. 235.

15. C(2025) 5053, Richtsnoeren van de Commissie betreffende de definitie van een artificiële-intelligentiesysteem zoals vastgesteld bij Verordening (EU) 2024/1689 (AI-verordening) ('Richtsnoeren'). Het European Law Institute (2024) heeft een kritische reactie gepubliceerd op de definitie

van AI-systeem in de verordening.

16. Zie ook overwegingen 22-23 Richtsnoeren.

17. J. Bijlsma, F.J. Bex & G. Meynen, 'Artificiële intelligentie en risicotaxatie. Drie kernvragen voor strafrechtjuristen', *NJB* 2019/2778, afl. 44, p. 3313-3319; G.M. de Vries e.a., 'AI-risicotaxatie: nieuwe kansen en risico's voor statistische voorspellingen van recidive', *Strafblad* 2021, afl. 2, p. 58-66.

18. *Risicovol algoritmegebruik*, p. 49. Zie ook de tabel waarin de hoor en wederhoor van de reclassering door de inspectie is gepubliceerd, onder 56 ('Hoor-wederhoor reactie Reclassering', www.inspectie-jenv.nl).

19. European Law Institute 2024, p. 16:

'This would imply that, to fall under the definition, an AI system only needs the "ability to generate outputs". Since almost all IT systems can generate outputs, this interpretation would vastly expand the scope of the AI Act.'

20. Overweging 12 AI-verordening.

21. S. Fazel e.a., 'Prediction of violent reoffending in prisoners and individuals on probation: a Dutch validation study (OxRec)', *Scientific reports* 2019, DOI:10.1038/s41598-018-37539-x.

22. M. Maas, E. Legters & S. Fazel, 'Professionaal risicotaxatie-instrument hand in hand. Hoe de reclassering risico's inschat', *NJB* 2020/1814 afl. 28, p. 2055-259. Zou geen sprake zijn van menselijke tussenkomst, dan zou het een verboden praktijk

zijn (art. 5 lid 1 onder d AI-verordening).

23. Op basis van een eerdere definitie in een conceptversie van de AI-verordening concludeerde Van Dijk eveneens dat *OxRec* binnen het bereik van de AI-verordening valt, evenals bijvoorbeeld de STATIC-99 (G. van Dijk, 'Predicting recidivism meets AI Act', *European Journal on Criminal Policy and Research* 2022, p. 407-423).

24. D.R. Cox, 'Regression models and life-tables', *Journal of the Royal Statistical Society. Series B (Methodological)* 1972, p. 187-220.

25. *Risicovol algoritmegebruik*, p. 49-50.

De inspectie verwijst in haar rapport naar de overweging 33 Richtsnoeren.

26. Zie paragraaf 3.1 van dit artikel.

27. Overweging 49-51 Richtsnoeren.

risicofactoren en de kans op recidive. Deze recidivekans kan niet worden benaderd door, bijvoorbeeld, het gemiddelde risico van een vergelijkbare groep te nemen. Hoewel *Cox proportional hazards*-regressie dus relatief simpel is vergeleken met moderne *machine learning*-algoritmen zoals neurale netwerken, kan de output van *OxRec* niet benaderd worden met een simpele vuistregel.

Het tweede argument is belangrijker en houdt verband met een mogelijk technisch onderscheid tussen statistiek en *machine learning* dat juridisch relevant is. Vaak wordt gezegd dat de statistische modellen van actuariële risicotaxatie minder complex zijn dan *machine learning*-modellen.²⁸ In het geval van *OxRec* is dat zeker waar. Bij klassieke statistische modellen zoals *Cox proportional hazards*-regressie bepaalt de mens vooraf de structuur. Dat wil zeggen: de mens bepaalt welke variabelen worden opgenomen in het model en hoe zij zich tot elkaar verhouden. Zo wordt vooraf bepaald dat leeftijd en aantal eerdere veroordelingen elk afzonderlijk meewegen. Het algoritme bepaalt vervolgens alleen het *gewicht* dat bij het bepalen van het recidiverisico aan deze variabelen moet worden toegekend. Bij complexere modellen die doorgaans als *machine learning* worden aangeduid – zoals neurale netwerken – doet het algoritme beduidend méér: het ontdekt zelf, bijvoorbeeld, dat jongere verdachten met veel veroordelingen een aparte risicogroep vormen, zonder dat de onderzoeker die combinatie vooraf heeft gespecificeerd. Het bepaalt dus zowel welke verbanden bestaan als hoe sterk zij zijn.

Het is echter niet de complexiteit van het algoritme op zichzelf dat bepaalt of sprake is van statistiek of *machine learning*. Modellen van vergelijkbare complexiteit als *Cox proportional hazards*-regressie, bijvoorbeeld het *naïve Bayes*-algoritme dat vaak voor spamfilters wordt gebruikt, worden in het algemeen als *machine learning* aangeduid,²⁹ terwijl ook deze uitsluitend gewichten binnen een vooraf bepaalde structuur schatten. Of een algoritme als statistiek of als *machine learning* wordt bestempeld, lijkt dus in ieder geval ten dele af te hangen van de discipline waarin de methode wordt toegepast en het doel waarvoor het wordt ingezet. Waar bijvoorbeeld criminologen en (forensisch-) gedragswetenschappers spreken van statistiek, kan een vergelijkbare methode in de computer- en datawetenschappen worden bestempeld als *machine learning*.³⁰ In de methodologische literatuur wordt het onderscheid dan ook primair gemaakt op basis van het doel: de klassieke statistiek richt zich op het gebruik van modellen en algoritmen voor het toetsen van hypothesen en het verklaren van verbanden,

Of een algoritme als statistiek of als *machine learning* wordt bestempeld, lijkt in ieder geval ten dele af te hangen van de discipline waarin de methode wordt toegepast en het doel waarvoor het wordt ingezet

AI en *data science* op het voorspellen van uitkomsten in nieuwe gevallen.³¹ *OxRec* is ontwikkeld en ingezet met het oog op dat tweede doel: de voorspelling van recidiverisico bij individuele personen. Daarmee sluit het aan bij de definitie van AI-systeem in de verordening, die als kenmerkend voorbeeld van output juist *voorspellingen* noemt, ongeacht of de onderliggende methode als statistiek of als *machine learning* wordt aangemerkt.

De vraag is welke conclusie moet worden getrokken uit de vaststelling dat geen principieel onderscheid bestaat tussen de algoritmes die worden gebruikt voor statistiek en *machine learning* en dat beide meer of minder complex kunnen zijn. In de richtsnoeren lijkt de Europese Commissie te kiezen voor een betrekkelijk ruime interpretatie van inferentievermogen, waardoor zowel statistische modellen als relatief eenvoudige *machine learning*-technieken onder de definitie van AI-systeem kunnen worden geschaard.³² De conclusie van de inspectie dat *OxRec* over inferentievermogen beschikt is in belangrijke mate gebaseerd op deze richtsnoeren. De richtsnoeren zijn echter niet bindend, niet altijd even duidelijk en bovendien aan verandering onderhevig. Een meer beperkte interpretatie is ook verdedigbaar. Indien hogere eisen worden gesteld aan het inferentievermogen, kan het zo zijn dat statistische modellen zoals *OxRec* én weinig complexe *machine learning*-methodes buiten de definitie van AI-systeem vallen.³³ De uitleg van het begrip inferentievermogen is uiteindelijk aan het Hof van Justitie.

3.3. Beschikken alle statistische risicotaxatie-instrumenten over inferentievermogen?

Er zijn dus argumenten om te denken dat de reikwijdte van de AI-verordening beperkter is dan de inspectie aanneemt. Tegelijkertijd zijn er ook redenen om te denken dat deze ruimer is. Zoals besproken, beschikt *OxRec* volgens de inspectie wel over inferentievermogen, maar andere door de reclassering gebruikte risicotaxatie-instrumenten niet. Hoewel *STATIC-99R*, *STABLE-2007* en *ACUTE-2007* ook gebaseerd zijn op statistisch onderzoek, zijn ze volgens de inspectie geen AI-systemen omdat de 'rekenregels in de *STATIC-99R*, *STABLE-2007* en *ACUTE-2007* door de mens bedacht zijn' en daarom geen inferentievermogen vertonen.³⁴ De inspectie baseert zich hier waarschijnlijk op de uitzondering in overweging 12 van de AI-verordening voor systemen die 'gebaseerd zijn op regels die uitsluitend door natuurlijke personen zijn vastgesteld om automatisch handelingen uit te voeren'. Deze uitzondering lijkt echter te zijn bedoeld voor eenvoudige regeluitvoerende systemen zoals database-systemen waarmee door middel van regels kan worden gesorteerd of gefilterd op basis van specifieke criteria (bijvoorbeeld 'zoek alle klanten die afgelopen maand een specifiek product hebben gekocht').³⁵ Maar de *STATIC-99R*, *STABLE-2007* en *ACUTE-2007* werken niet op die manier. De uitzondering zou daarom niet van toepassing kunnen zijn.

Het argument dat de inspectie gebruikt om *OxRec* te onderscheiden van andere risicotaxatie-instrumenten is op zichzelf dan ook niet overtuigend. En omdat geen principieel verschil bestaat tussen eenvoudige *machine learning* en statistische modellen rijst ook ten aanzien van deze andere instrumenten de vraag in hoeverre sprake is van inferentievermogen. Ook hier is het antwoord afhankelijk van hoe

Het concrete voorbeeld van *OxRec* illustreert niet alleen een afbakeningsprobleem, maar brengt ook een spanning aan het licht die in de kern van de verordening zit

ruim of hoe beperkt het begrip inferentievermogen wordt geïnterpreteerd.³⁶ Een voorwaarde is natuurlijk wel dat aan de andere elementen van de definitie van AI-systeem is voldaan. Sommige statistische risicotaxatie-instrumenten kunnen ook op papier worden ingevuld. In dat geval is geen sprake van een 'op een machine gebaseerd systeem', zoals de definitie van AI-systeem vereist.

4. Conclusie

De discussie over *OxRec* legt een fundamenteel dilemma bloot waar de definitie van AI-systeem in de AI-verordening ons voor stelt. Wordt inferentievermogen ruim geïnterpreteerd, zoals de richtsnoeren suggereren en de inspectie doet, dan dreigt de verordening in de praktijk een statistiek- of algoritmeverordening te worden. Klassieke (statistische) methoden en algoritmen worden breed toegepast in onder andere het onderwijs, de financiële dienstverlening en de geneeskunde. Het aanmerken van *OxRec* als AI-systeem trekt een grens die, consequent doorgevoerd, een groot deel van die alledaagse praktijk potentieel binnen het bereik van de verordening brengt, met mogelijk verstrekkende gevolgen die vermoedelijk niet werden beoogd. Bij een enge interpretatie van inferentievermogen dreigt een tegenovergesteld probleem. In dat geval vallen algoritmen die op basis van statistische analyse een risicoscore genereren die de vrijheid van een persoon kan beïnvloeden buiten de verordening. De verordening zou dan geen bescherming bieden in de context – risicotaxatie in de strafrechtspleging – die expliciet als hoogrisico is aangemerkt. Overigens moet bedacht worden dat de andere elementen uit de definitie van AI-systeem, bijvoorbeeld dat het systeem een zekere mate van autonomie moet hebben en invloed op de omgeving moet kunnen uitoefenen, ook bijdragen aan de afbakening van het toepassingsbereik van de AI-verordening.

Het concrete voorbeeld van *OxRec* illustreert niet alleen een afbakeningsprobleem, maar brengt ook een spanning aan het licht die in de kern van de verordening zit: de definitie moet rechtszekerheid bieden én flexibel genoeg zijn om nieuwe technologie te omvatten, maar schiet op dit moment tekort in beide opzichten. Die onzekerheid klinkt niet door in het betrekkelijk stellige oordeel van de inspectie dat *OxRec* wel een AI-systeem is en drie andere risicotaxatie-instrumenten niet. Het Hof van Justitie zal hier uiteindelijk knopen moeten doorhakken.

De gevolgen van een ruime of beperkte interpretatie van inferentievermogen voor andere domeinen dan de risicotaxatie kunnen wij niet overzien. Die gevolgen moeten wel een rol spelen in het antwoord op de vraag hoe inferentievermogen moet worden geïnterpreteerd. Het door ons gesignaleerde dilemma kunnen wij daarom niet oplossen. In dit opzicht bestaat vermoedelijk een zekere spanning tussen de doelstelling van de AI-verordening om de rechtsstaat te beschermen en het ondersteunen van innovatie.

Ter afsluiting wijzen wij op het risico dat interpretaties van onder meer inferentievermogen gaan divergeren, vooral waar niet evident sprake is van zelflerende algoritmes. Het risico op uiteenlopende interpretaties van inferentievermogen wordt versterkt doordat AI in vrijwel alle maatschappelijke domeinen wordt gebruikt, waardoor zij raakt aan de taken van heel diverse (toezichhoudende) instanties – en dat in een unie die bestaat uit 27 lidstaten. Daardoor lijkt een reëel risico op een gebrek aan rechtszekerheid te ontstaan. Het concrete voorbeeld van de discussie over *OxRec* kan dienen om op nationaal en Europees niveau tot een nadere afbakening van het begrip AI-systeem te komen. Als de Autoriteit Persoonsgegevens inderdaad wordt aangemerkt als markttoezichtautoriteit, dan is daar voor haar een belangrijke taak weggelegd. •

28. Zie nader over het onderscheid tussen actuariële risicotaxatie en risicotaxatie met behulp van AI: G.M. de Vries e.a. 2022. Zie ook J. Bijlsma, F.J. Bex & G. Meynen, 'Artificiële intelligentie en risicotaxatie. Drie kernvragen voor strafrechtjuristen', *NJB* 2019/2778, afl. 44, p. 3313-3319
29. M.I. Jordan & T.M. Mitchell, 'Machine learning: trends, perspectives, and prospects', *Science* 2015, p. 255-260.
30. Illustratief is dat T. Hastie, R. Tibshirani & H. Friedman in hun standaardwerk *The elements of statistical learning* (New York: Springer 2009) simpele statistische metho-

den vergelijkbaar met *Cox proportional hazard-regressie* en complexere *machine learning*-methoden zoals neurale netwerken onder één noemer – *statistical learning* – behandelen.
31. L. Breiman, 'Statistical modeling: The two cultures', *Statistical Science* 2001, 16, p. 199-231; D. Bzdok, N. Altman & M. Krzywinski, 'Statistics versus machine learning', *Nature Methods* 2018, p. 233-234.
32. Op de vraag van europarlementariërs Voss en Hahn van 18 november 2024 (E-002564/2024) hoe de Europese Commissie wil bewerkstelligen dat eenvoudige

statistische technieken buiten de definitie van 'AI-systeem' vallen, is namens de Commissie geantwoord: 'These guidelines explain that neither Article 3(1) AI Act, nor Recital 12 AI Act which clarifies that provision, explicitly excludes statistical methods from the scope of the definition of "AI system".'
33. Smuha & Young 2025, p. 235: 'The definition's emphasis on the system making inferences seems to exclude more traditional or rule-based AI systems from its scope, despite their significant potential for harm. Ultimately, it will be up to the courts to

decide how this definition must be interpreted in case of a dispute.'
34. *Risicovol algoritmegebruik*, p. 27.
35. Overweging 46 Richtsnoeren.
36. Vgl. Van Dijck 2022, p. 409. Een mogelijkheid is dat deze modellen om een andere reden moeten worden aangemerkt als AI, namelijk als de 'op logica en kennis gebaseerde benadering' van AI zoals die in overweging 39 van de richtsnoeren worden genoemd, hoewel ook sprake is van verschillen met deze 'expertsystemen'. De inspectie bespreekt die mogelijkheid niet in haar rapport.