

# 21. De ggz-behandelaar in het Wilde Westen van de Natural Language Processing Tools

## De mogelijkheden en risico's van het gebruik van AI in de spreekkamer

ROLAND BERTENS & FLORIANE JASPERS

Kunstmatige intelligentie (AI) heeft binnen korte tijd grote sprongen vooruit gemaakt, zowel in mogelijkheden en toepassingen als in acceptatie en gebruik. Voor de EU was dat aanleiding om regelgeving op te stellen: de **Artificiële Intelligentie (AI) Verordening**. In de gezondheidszorg lijkt het in het bijzonder van belang om rekening te houden met de mogelijkheden maar ook risico's die het gebruik van AI mee kan brengen voor de fundamentele (grond)rechten van patiënten. Patiënten behoren immers tot een kwetsbare groep burgers.

### Inleiding

De definitieve versie van de Artificiële Intelligentie (AI) Verordening is op 9 december 2023 aangenomen in Europa<sup>1</sup>. Deze verordening, die na inwerkingtreding dwingend recht zal zijn in alle Europese lidstaten, markeert een belangrijke stap in de regulering van het gebruik van kunstmatige intelligentie, algoritmes en *machine learning* binnen Europa.<sup>2</sup> Het doel van de verordening is om een balans te zoeken tussen de voordelen en nadelen die kunstmatige intelligentie kan bieden in allerlei economische en maatschappelijke sectoren, waaronder de gezondheidszorg. De verordening is er mede op gericht de risico's die AI met zich mee kan brengen, onder andere op het gebied van profilering, social scoring, discriminatie en meer, te beperken. Het gaat erom innovatie in AI te ondersteunen en tegelijkertijd een hoog niveau van bescherming van de gezondheid, de veiligheid en grondrechten te waarborgen, met inbegrip van de democratie, de rechtsstaat en de bescherming van het milieu.<sup>3</sup>

1 Voluit gaat het om de Verordening tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie. De meest recente versie van de verordening is te vinden via: <https://artificialintelligenceact.eu/>. Ten behoeve van dit artikel is gebruikgemaakt van de versie van 16 april 2024, waarin nog kleine redactionele aanpassingen zijn gedaan (<https://artificialintelligenceact.eu/the-act/>). Qua terminologie bedoelen wij met 'artificiële intelligentie' (kortweg AI) hetzelfde als met 'kunstmatige intelligentie'.

2 De daadwerkelijke inwerkingtreding zal gefaseerd verlopen, maar de meeste bepalingen uit de Verordening zullen 24 maanden na publicatie in het *Publicatieblad van de Europese Unie* van kracht worden. Publicatie wordt nu verwacht tussen mei en juli 2024.

3 Art. 1 lid 1 AI Verordening. Zie ook overwegingen 4 t/m 8 bij de Verordening.

In de gezondheidszorg lijkt het in het bijzonder van belang om rekening te houden met de mogelijkheden maar ook risico's die het gebruik van AI mee kan brengen voor de fundamentele (grond)rechten van patiënten. Patiënten behoren immers tot een kwetsbare groep burgers. Zij wenden zich wegens (geestes)ziekte tot een hulpverlener, die vrijwel altijd in een sterkere machts- en kennispositie verkeert. Zaken als het goed informeren van de patiënt en het verkrijgen van toestemming voor een behandeling zijn daarom ook van essentieel belang.

Dat roept vragen op over het gewenste gebruik van AI in de (geestelijke) gezondheidszorg. Sinds de lancering van chatbot ChatGPT in november 2022 wordt er binnen de hedendaagse praktijk volop geëxperimenteerd met verschillende artificiële intelligentie-tools. Voor zij die deze tools testen, wordt de potentie om behandlungsprocessen aanzienlijk te transformeren snel duidelijk. Desalniettemin roepen deze applicaties en het gebruik daarvan ethische en juridische vragen op. Wat moet er verstaan worden onder AI? Hoe mogen hulpverleners gebruikmaken van deze tools (en hoe niet)? En hoe verhouden de nieuwe regels rondom AI zich tot de klassieke patiëntenrechten?

In deze bijdrage gaan wij in op enkele van deze vraagstukken. Wij leggen de focus daarbij op het gebruik van 'Natural Language Processing Tools' (zoals ChatGPT) in de dagelijkse praktijk van de behandelaar in de (j)ggz. Daarbij bespreken wij ook een casus die uit de praktijk gegrepen is, namelijk het gebruik van NLP voor dossiervoering.

### Natural Language Processing: revolutionaire (?) AI

NLP staat voor Natural Language Processing, oftewel natuurlijke taalverwerking. Tools zoals ChatGPT, Gemini,

RoBERTa, Jasper, Bing enzovoort maken een integraal deel uit van het snel uitdijende veld van NLP. Het is een onderdeel van kunstmatige intelligentie en computerwetenschappen dat zich bezighoudt met de interactie tussen computers en menselijke (natuurlijke) taal. Hoewel er momenteel veel aandacht voor is, is dit geen nieuwe wetenschap: ELIZA was een vroeg computerprogramma voor natuurlijke taalverwerking, ontwikkeld tussen 1964 en 1967 aan het MIT door Joseph Weizenbaum.

Specifieke functies van NLP's zijn onder andere het vertalen van tekst, spraakherkenning, tekstuele sentimentanalyse, chatbots en virtuele assistenten, het zoeken en ophalen van informatie uit grote databanken op basis van gebruikersvragen en het categoriseren en organiseren van tekst in relevante thema's of categorieën. Met name deze laatste functies hebben een grote aantrekkingskracht op de ggz-professional: een NLP is – technisch – in staat een dossier op te bouwen, differentiaaldiagnoses te stellen en een behandelvoorstel te genereren aan de hand van gesprekken tussen de patiënt en de professional.

NLP-modellen zijn opgebouwd met talrijke besluitvormingslagen en miljoenen parameters, waarvan de onderlinge interacties moeilijk te doorgronden zijn. Dit maakt het traceren van besluitvormingsprocessen binnen het model uitdagend tot onmogelijk, zeker voor het merendeel van behandelaren en patiënten. De zelflerende aard van NLP-systemen, die zonder menselijke sturing relevante kenmerken uit data extraheren, versterkt hun ondoorzichtige natuur.

Binnen de kaders van de nieuwe Europese AI Verordening, wordt er een risicogebaseerde benadering gehanteerd ten aanzien van kunstmatige-intelligentiesystemen. Daarbij worden AI-systemen ingedeeld in vier verschillende risiconiveaus: onaanvaardbaar risico, hoog risico, beperkt risico en minimaal risico. Deze risicobenadering is functioneel: afhankelijk van wat een AI-systeem precies doet, komt het in een van de vier categorieën terecht. Art. 6 Verordening bevat daarbij de harde bovengrens: daarin wordt een aantal typen AI-praktijken verboden binnen de Europese markt wegens het onaanvaardbare risico dat zij creëren. Zo is het verboden om een AI-systeem in de handel te brengen dat subliminale technieken of doelbewust misleidende technieken gebruikt, waarmee mensen worden bewogen om een besluit te nemen dat ze anders niet hadden genomen en dat waarschijnlijk tot aanzienlijke schade zal leiden.<sup>4</sup> Ook zijn AI-toepassingen verboden die gebruikmaken van de kwetsbaarheden van een persoon als gevolg van hun leeftijd, handicap of een specifieke sociale of economische omstandigheid, om het gedrag van die persoon 'wezenlijk te verstoren' en indien die verstoring vervolgens leidt tot aanzienlijke schade.<sup>5</sup> En er zijn nog meer verboden applicaties.

In het geval van een NLP-systeem is niet zo snel in te zien hoe dat bewust gebruik zou maken van subliminale beïnvloeding of manipulatie. Het is echter wel voorstelbaar

dat er, binnen het kader van de gezondheidszorg, mogelijk specifieke chatprogramma's zullen worden ontwikkeld die gericht gaan zijn op mensen met een (mentale) aandoening. Als dat het geval is, dan zal een producent ervoor moeten waken dat het gedrag van deze personen niet 'wezenlijk verstoord' wordt door gebruik van het programma. Over wat er precies onder iets als 'wezenlijke verstoring' moet worden verstaan, zullen na inwerkingtreding van de AI Verordening ongetwijfeld discussie en jurisprudentie ontstaan. Maar het is dus duidelijk dat de Brusselse wetgever paal en perk heeft willen stellen aan bepaalde typen inzet van AI.

Praktisch gesproken zal de categorie van de 'hoogrisico' applicaties in de gezondheidszorg echter een belangrijker rol gaan spelen. Welke AI's er 'hoogrisico' zijn, is in de AI Verordening zelf niet opgenomen. In plaats daarvan moet worden vastgesteld wat als hoogrisicosystemen geldt door aan te sluiten bij andere Europese verordeningen die worden genoemd in Bijlage I bij de AI Verordening (Europese wetgeving staat doorgaans niet bekend om haar eenvoud en deze verordening is daarop helaas geen uitzondering). In die bijlage wordt verwezen naar de Verordening Medische Hulpmiddelen uit 2017. Kort samengevat is het zo dat, indien een AI-(software)applicatie ten doel heeft om te worden gebruikt voor onder meer diagnosestelling, prognosticering, behandeling of verlichting van ziekte, letsel of beperking, deze al snel zal vallen onder de definitie van 'medisch hulpmiddel' in de zin van de Verordening Medische Hulpmiddelen.<sup>6</sup> En voor zover dat 'hulpmiddel' vervolgens gecertificeerd moet worden door een 'notified body' in de zin van de Verordening Medische Hulpmiddelen, is die AI daarmee automatisch ook een hoogrisicoapplicatie in de zin van de AI Verordening.<sup>7</sup> Deze classificatie vereist vervolgens van ontwikkelaars en gebruikers van AI in de ggz *compliance* met strikte eisen op het gebied van transparantie, logging, toezicht en de mogelijkheden van menselijke interventie bij gebruik van de AI.<sup>8</sup>

### NLP als tool in de ggz: medische en juridische risico's

AI-applicaties die worden ontwikkeld voor de (j)ggz zullen op basis van het voorgaande al heel snel onder het hoogrisicodomein vallen, voor zover zij ten doel hebben daadwerkelijk te worden ingezet bij de diagnosticering en behandeling van patiënten. Het opmerkelijke is nu natuurlijk dat veel NLP's niet met *dat doel* zijn gebouwd. Een programma als ChatGPT maar ook Gemini van Google is in staat om op basis van tekstininput tekstuele output te leveren. Aldus kan zo'n programma worden gebruikt om vragen te stellen over medische diagnoses. Maar je kunt het in theorie ook de belastingaangifte laten doen. Dat maakt dat zulke

4 Art. 5 lid 1 sub a AI Verordening.

5 Art. 5 lid 1 sub b AI Verordening.

6 Zie de definitie van 'medisch hulpmiddel' in art. 2 lid 1 Verordening (EU) 2017/745.

7 Zie art. 6 lid 1, AI Verordening in samenhang met Bijlage I, onder 11, van de AI Verordening.

8 Zie onder meer art. 8 t/m 15 van de AI Verordening.

programma's – ook na inwerkingtreding van de AI Verordening – lang niet altijd als 'hoogrisico' zullen worden gekwalificeerd (laat staan als medisch hulpmiddel).<sup>9</sup> De verleiding zal dan ook groot zijn om te denken dat, nu die programma's kennelijk zomaar op de markt kunnen worden gebracht, er niets mis is met het gebruik in de patiëntenzorg. Wij zien dan ook dat er inmiddels al serieus wordt geëxperimenteerd met het gebruik van NLP in de patiëntenzorg. Wij signaleren daarbij twee risico's.

Het is ten eerste van belang om te allen tijde in gedachten te houden dat, zelfs als een NLP (ook na inwerkingtreding van de AI Verordening) níét wordt geclassificeerd als een hoogrisicoapplicatie, deze toch kan worden ingezet in de gezondheidszorg en dat serieuze consequenties kan hebben voor zoets als diagnosticering. Een voor de hand liggend voorbeeld is de differentiaalanamnese: de zorgverlener kan het ziektebeeld van de patiënt (symptomen, anamnese of diagnose) invoeren in een NLP en vervolgens de vraag stellen welke diagnoses daar mogelijk bij horen. De NLP put daarbij uit een aanzienlijke hoeveelheid erkende medische teksten om tot een antwoord te komen. In het beste geval leidt dat tot een teveel aan mogelijke diagnoses (met als gevolg wellicht de inzet van onnodige onderzoeken en therapieën). In het slechtste geval mist een AI echter een cruciale diagnose. Het is in dit voorbeeld duidelijk dat een NLP die niet direct is gebouwd voor medische handelingen en daar toch voor wordt ingezet serieuze risico's voor een patiënt kan creëren. Dat is anders als er straks medische AI's op de Europese markt komen die uit en te na getest en gecertificeerd zijn. Bij zulke programma's moet de producent vooraf voldoen aan allerlei belangrijke transparantie- en veiligheidscriteria. Vooralsnog bevinden wij ons echter in het 'Wilde Westen' – en dat zal mogelijk altijd het geval blijven bij het gebruik van 'general purpose AI' of NLP's die niet specifiek voor de zorg zijn ontwikkeld.

Ten tweede signaleren wij een juridisch risico. Met de komst van de AI Verordening zullen sommige hulpverleners denken (en misschien wel hopen) dat daarmee de Nederlandse wetgeving aangaande patiëntenrechten terzijde kan worden geschoven. Dat is echter niet zo. De verordening zal, anders dan een Europese richtlijn, vanaf inwerkingtreding direct dwingend recht zijn, ook in Nederland. Maar zij komt daarmee náást geldende nationale wetgeving te staan (en is daarnaast ook afgestemd op andere Europese wetgeving, zoals de AVG). Dat heeft tot gevolg dat het gebruik van AI in de zorg vorm zal moeten krijgen binnen de kaders zoals wij die in Nederland hebben gesteld voor de patiëntenzorg. Daarbij gaat de WGBO voorop.

De AI Verordening zal weliswaar nieuwe rechten creëren voor personen die gebruikmaken van AI's. Een voorbeeld is art. 51 lid 1 AI Verordening. Dat bepaalt dat de maker van een AI-systeem dat bedoeld is om direct te interageren met natuurlijke personen, die AI zo ontwerpt dat het voor de gebruiker direct duidelijk is dat hij te maken heeft met een AI-systeem (tenzij dat echt evident is). Maar stel: een

NLP wordt gebruikt zoals in het voorbeeld hiervoor, namelijk door een hulpverlener die het de differentiaaldiagnose (mede) door de AI laat stellen. De hulpverlener is in dat geval de persoon die interageert met de NLP. Hij moet op basis van art. 51 AI Verordening dus weten dat hij te maken heeft met een AI. Maar als hij vervolgens niet aan de patiënt vertelt dat hij gebruik heeft gemaakt van een AI om tot zijn diagnose te komen, is het duidelijk dat de patiënt niet zoveel heeft aan het recht uit art. 51. De positie van de patiënt en de verplichtingen van de hulpverlener moeten dan ook worden ingekleurd op basis van de patiëntenrechten zoals die in de WGBO staan.

### Casus: het gebruik van NLP bij de dossiervoering

Om dit nader te illustreren, bespreken wij een voorbeeld dat uit de praktijk is gegrepen: het gebruik van NLP om dossiers te voeren. Dat kan in allerlei vormen: professionals in de ggz kunnen consulten opnemen en ze vervolgens laten optekenen en uitwerken door een NLP, bijvoorbeeld ChatGPT. De intake wordt uitgewerkt in een beschrijvende diagnose, er wordt een huisartsenbrief gegenereerd en een behandelplan opgesteld. Automatisch wordt er een bericht verstuurd naar de eerdere betrokkenen en informatie opgevraagd die voor de behandelaar wordt samengevat. Tevens wordt er een bloedonderzoek aangevraagd.

De (administratieve) voordelen van zulk gebruik zijn evident: welke ggz-professional wil er nou niet minder tijd kwijt zijn met het beheren van dossiers? Bezien we deze casus echter juridisch, dan zitten daar de volgende haken en ogen aan.

Ten eerste is hier sprake van een evident privacy- en beroepsgeheimrisico. Op Europeesrechtelijk niveau is er een grondslag voor verwerking van bijzondere (medische) persoonsgegevens, zónder dat sprake hoeft te zijn van expliciete toestemming.<sup>10</sup> Echter, zodra gegevens gedeeld worden met anderen dan de patiënt zonder diens toestemming, is dat direct in strijd met Nederlandse wetgeving, namelijk art. 7:457 lid 1 Burgerlijk Wetboek (de WGBO).<sup>11</sup> Dat het invoeren van gegevens van een patiënt, zonder diens toestemming, in een NLP waarvan doorgaans niet helder is waar die gegevens terechtkomen of hoe die verwerkt worden, op gespannen voet staat met het beroepsgeheim hoeft dan ook geen uitgebreid betoog.

En hierbij wordt met patiëntgegevens ook, niet, zoals in sommige essays of artikelen de laatste tijd wordt gesuggereerd, enkel de geboortedatum en achternaam bedoeld. Gezien het feit dat NLP's op basis van kenmerkextractie functioneren (automatisch de relevante informatie of 'kenmerken' uit ruwe data identificeren en extraheren om deze te gebruiken voor verdere analyse, lering of model-lering) is volstrekt onduidelijk of bijvoorbeeld gegevens

9 Hoewel voor AI's die 'systemic risk' creëren dan wel weer specifieke eisen gelden volgens art. 55 AI Verordening.

10 Art. 9 lid 2 sub h Algemene Verordening Gegevensbescherming, nader uitgewerkt in art. 30 lid 3 sub a Uitvoeringswet Algemene Verordening Gegevensbescherming (UAVG).

11 En overigens ook met art. 30 lid 4 UAVG.

uit de (sociale) anamnese niet herleidbaar kunnen zijn. De term ‘anonimisering’ heeft in discussies over de Algemene Verordening Gegevensbescherming al aan kracht ingeboet, omdat inmiddels duidelijk is dat data echt ‘anoniem’ maken steeds moeilijker is geworden. Met die voorzichtigheid moet ook worden aangekeken tegen het invoeren van dossierdata in NLP.

Daarnaast valt er vanuit het perspectief van ‘informed consent’ (art. 7:448 jo. art. 7:450 BW) het nodige aan te merken op het gebruik van NLP en dossiervoering. Als een NLP enkel gebruikt wordt als een ‘schrijfassistent’ zal geen sprake zijn van invloed op de behandeling. Maar zodra een hulpverlener zich begint te verlaten op koppelingen, aanbevelingen, analyses of beslismogelijkheden die (subtiel) worden gedaan door een NLP bij verwerking van het dossier, rijst de vraag of de patiënt geïnformeerd zal moeten worden dat gebruik is gemaakt van zo’n programma. De letter én geest van de wet wijzen dan al snel in die richting. De WGBO gaat namelijk uit van het optimaal beschermen van de patiënt en bewaken van diens autonomie, onder meer door de ondergeschikte kennis- en machtspositie die hij heeft ten opzichte van zijn hulpverlener wat te mitigeren. Zo is in 2020 ook ‘shared-decision making’ explicieter in de WGBO opgenomen. Het is niet waarschijnlijk dat de wetgever daarbij heeft bedoeld dat AI’s ook deel kunnen nemen aan zulke decisionmaking zónder dat de patiënt daarvan op de hoogte wordt gesteld.

Het voorgaande betekent niet dat het gebruik van NLP’s in of ten behoeve van de patiëntenzorg verboden is of zou moeten zijn. Sterker nog: wij juichen zulk gebruik toe, voor zover dat hulpverleners kan ontlasten bij administratieve werkzaamheden en het voordelen oplevert voor de patiënt. Maar het is wel van belang om zeer scherp te zijn op de wijze waarop NLP’s worden ingezet, hoe de data waarmee ze worden ‘gevoed’ worden verwerkt en waar deze terecht komt, en in hoeverre patiënten op de hoogte worden gesteld van het gebruik van dit soort nieuwe applicaties.

## Conclusie

Ondanks het optimisme (wat wij overigens delen) van veel collega’s over de inzet van NLP in de ggz is het cruciaal om rekening te houden met deze vrij heldere wettelijke kaders. Het invoeren van medische gegevens of delen van medische dossiers in vrij verkrijgbare AI-systemen is onder de huidige wetgeving niet zonder meer toegestaan. Voor zover dat toch gedaan wordt, is het goed om daarbij – conform de letter en geest van de wet – patiënten optimaal te betrekken. Daarnaast is het de taak van zorgverleners om de privacy van patiënten te beschermen en daarmee te waarborgen dat patiënten ook het vertrouwen behouden in de zorg die zij ontvangen.

In de praktijk betekent dit dat behandelaren in de ggz zich bewust moeten zijn van de juridische kaders die het gebruik van NLP-tools omgeven. Het vereist dat zij op de hoogte blijven van de laatste ontwikkelingen op het gebied van wetgeving en ethische richtlijnen en dat zij nauw samenwerken met juridisch en ethisch adviseurs om te zorgen voor compliantie. Het vraagt ook om een voortdurende dialoog met patiënten over het gebruik van deze technologieën, inclusief transparantie over de voordelen, risico’s en de mate van betrokkenheid van AI in hun behandeling.

### Over de auteurs

#### Roland Bertens

Advocaat Zorg en Sociaal Domein bij Van Benthem & Keulen  
Advocaten en daarnaast universitair docent gezondheidsrecht bij het UMC Utrecht.

#### Floriane Jaspers

Kinder- en Jeugdpsychiater i.o. bij Levvel. Ze doet onderzoek naar Conversational Artificial Intelligence voor het Amsterdam UMC en is mede-initiatiefnemer van het kennisplatform AI in de GGZ.  
Mocht u interesse hebben om met Roland en Floriane samen te werken, uw kennis te delen of bent u op zoek naar meer informatie, neem dan contact op via het kennisplatform AI in de GGZ.