



Universiteit
Leiden
The Netherlands

Computer says N0? Over zorgvuldigheid en ongeschikte hulpzaken bij gebruik van medische AI-output

Staalduinen, J.H. van

Citation

Staalduinen, J. H. van. (2025). Computer says N0?: Over zorgvuldigheid en ongeschikte hulpzaken bij gebruik van medische AI-output. *Tijdschrift Voor Gezondheidsrecht*, 49(3), 225-241.
doi:10.5553/TvGR/016508742025049003004

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/4292273>

Note: To cite this publication please use the final published version (if applicable).

ARTIKELEN

Computer says NO?

Over zorgvuldigheid en ongeschikte hulpzaken bij gebruik van medische AI-output*

Mr. drs. J.H. van Staalduinen**

1. Inleiding

Beslissingsondersteunende AI-systemen kunnen in de medische praktijk een nuttige bijdrage leveren door sneller en/of accurater informatie van de patiënt te verwerken en/of van duiding te voorzien. Denk bijvoorbeeld aan software die afwijkend weefsel, zoals tumorweefsel, detecteert ('NO' is een aanduiding binnen het TNM-tumorclassificatiesysteem¹). Ook als de hulpverlener de output vervolgens nog handmatig beoordeelt, kan het gebruik van deze software zorgen voor tijdswinst en vroegere detectie. Als en naarmate AI-systemen geavanceerder worden, is het op termijn echter niet uitgesloten dat de output van deze AI-systemen niet altijd handmatig meer te checken is, of dat dat wel mogelijk is, maar onredelijk veel tijd kost.²

Als het desondanks de moeite waard is dergelijke systemen te gebruiken, moet de hulpverlener in zekere mate kunnen vertrouwen op de output van zo'n systeem. Bij sommige apparaten die in het ziekenhuis gebruikt worden, geldt dit nu eigenlijk al. Meer nog dan bij handmatig geprogrammeerde software, kan het bij AI-systemen echter voorkomen dat er iets op onverwachte manier fout gaat in het systeem, ook als het systeem goed is onderhouden. Als de hulpverlener zich nietsvermoedend door verkeerde output laat beïnvloeden – ook als daarnaast nog andere informatiebronnen worden betrokken – kan er schade ontstaan bij de patiënt. De hulpverlener kan dan niet worden verweten onzorgvuldig te zijn geweest. In zo'n situatie zou de hulpverlener niettemin aansprakelijk kunnen zijn op grond van artikel 6:74

* Deze bijdrage is geschreven in het kader van het ZonMw-project 'DECIDE-VerA' (projectnr. 08540122120004).

** Jan van Staalduinen (ORCID 0000-0001-7937-9006) is PhD-fellow bij het Instituut voor Privaatrecht van de Universiteit Leiden en bereidt een proefschrift voor over civielrechtelijke aansprakelijkheid voor medische AI-systemen.

1 Deze aanduiding betekent: geen lymfekliermetastase.

2 C.A. Tschider, 'Humans Outside the Loop', *Yale Journal of Law & Technology* (26) 2024, afl. 2, p. 327-354 en p. 361-368.

jo. 6:77 BW, als vast komt te staan dat het systeem kwalificeert als hulpzaak;³ dat het op het moment van gebruik ongeschikt was; dat hierdoor tekort is geschoten in de nakoming van de verbintenissen uit hoofde van de geneeskundige behandelingsovereenkomst; dat hierdoor schade bij de patiënt is ontstaan; en dat het niet onredelijk is deze tekortkoming aan de hulpverlener toe te rekenen.

Het is mijns inziens niet nodig en zelfs contraproductief om een dergelijke tekortkoming 'alsnog' te duiden als onzorgvuldigheid bij de aanschaf, controle of toepassing van de hulpzaak teneinde tot een sluitende interpretatie van de jurisprudentie rond artikel 6:77 BW te komen. Ik zal dat verderop in dit artikel toelichten.

In de Nederlandse literatuur over de juridische aspecten van medische AI gaat onder andere aandacht uit naar de verplichtingen uit de AI Act:⁴ gebruiksinstructies naleven, adequaat opgeleid menselijk toezicht verzorgen, relevante data invoeren, de werking van het systeem monitoren en logbestanden bewaren.⁵ Uiteraard wordt verwacht dat alleen CE-gecertificeerde systemen worden gebruikt.⁶ Beperktere aandacht is er vooralsnog voor de centrale medisch-inhoudelijke activiteit: het interpreteren, kritisch beschouwen en wegen van de output van een AI-systeem bij het nemen van een concrete behandelbeslissing.⁷

Deze bijdrage gaat over de vraag hoe deze activiteit aansprakelijkheidsrechtelijk genormeerd is. Het gaat daarbij uiteindelijk om de vraag hoe bij het gebruik van een beslissingsondersteunend AI-systeem moet worden beoordeeld of er sprake is van een tekortkoming, waarbij opgehelderd moet worden in hoeverre dit toegeschreven kan worden aan onzorgvuldigheid van de hulpverlener en in hoeverre dit toegeschreven kan worden aan ongeschiktheid van het AI-systeem.

Deze bijdrage is als volgt gestructureerd. Eerst wordt besproken welke kenmerken van AI bijzondere invloed hebben op het medische besluitvormingsproces en welke zorgen dit kan opleveren voor het aansprakelijkheidsrecht (par. 2). Vervolgens

- 3 Hier wordt aangenomen dat art. 6:77 BW ook voor software kan gelden, hoewel dat strikt genomen nog geen uitgemaakte zaak is. Zie A.J. Zijlstra, 'Civiele aansprakelijkheid voor het gebruik van medische applicaties', *TvGR* 2016, p. 441-454, zie par. 4.2.1; M.R.L. Ruesen, 'Ongeschikte medische apps: wie is aansprakelijk?', *Computerrecht* 2017/49, par. 1.3; T.J. de Graaf & I.S. Wuisman, 'Contractuele aansprakelijkheid voor het gebruik van AI-systemen in het licht van Europese voorstellen', *WPNR* 2021, afl. 7348, par. 5; D.S. Kroese, J. Wassink & S.E. de Rey, 'Contractuele aansprakelijkheid voor ongeschikte hulpzaken en gebrekkige hulpvoorwerpen in Nederland en België', in: S.E. de Rey (red.), *Het nieuwe Belgische verbintenissenrecht vanuit Nederlands perspectief*, Zutphen: Uitgeverij Paris 2024, par. 4.3; H.B. van Kolschooten & R.P. Wijne, 'Diagnosticeren met AI en de aansprakelijkheid van hulpverleners', *NJB* 2024/1414, par. 4.2.
- 4 Verordening 2024/1689 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (AI Act) en de voorgestelde AI-aansprakelijkheidsrichtlijn (Voorstel voor een Richtlijn betreffende de aanpassing van de regels inzake niet-contractuele civielrechtelijke aansprakelijkheid aan artificiële intelligentie (AILD) (2022/0303/COD, COM(2022) 496 final)).
- 5 Art. 26 lid 1, 2, 4, 5 en 6 AI Act. Zie recent Van Kolschooten & Wijne 2024 en R. Jessen, 'Op het snijvlak van technologie, zorg en recht: aansprakelijkheid van het ziekenhuis bij gebruik van besluitvormingsalgoritmen', *TvGR* 2024, p. 33-43.
- 6 Van Kolschooten & Wijne 2024, p. 1768.
- 7 Beknopt hierover Van Kolschooten & Wijne 2024, p. 1768.

wordt een onderwerp centraal gesteld dat enkele jaren geleden aandacht kreeg naar aanleiding van de *PIP*- en *Miragelplombe*-arresten van de Hoge Raad:⁸ de verhouding tussen de plicht om goed hulpverlenerschap te betrachten en de plicht om geen ongeschikte hulpzaken te gebruiken. In deze bijdrage wordt verdedigd dat dit twee onderscheiden, maar samenhangende verplichtingen zijn, met zowel een eigen maatstaf voor tekortkoming als een eigen toerekeningskader (par. 3). De praktische vervolgvraag luidt hoe in een concreet geval dan onderscheid kan worden gemaakt tussen deze twee plichten, welke wordt uitgewerkt voor de situatie waarin beslissingsondersteunende AI is gebruikt (par. 4). Ten slotte wordt besproken in hoeverre de in paragraaf 2 gesignaleerde zorgen in het aldus ontwikkelde Nederlandse kader terecht zijn (par. 5), om af te sluiten met een conclusie (par. 6).

Ik beperk mij tot het materiële verbintenissenrecht en daarbinnen hoofdzakelijk tot de omvang van de op de hulpverlener rustende verbintenis(sen).⁹

2. Wat maakt AI bijzonder?

2.1 AI als technische en normatieve black box

Medische AI-adviesystemen ('clinical decision support systems') kunnen worden gezien als een variant van oudere expertsystemen. In tegenstelling tot deze eerdere systemen, zijn AI-adviesystemen niet meer geheel gebaseerd op voorgeprogrammeerde regels. AI-systemen 'leren' hun eigen regels uit een dataset tijdens een trainingsproces. Dit kan het lastiger maken om te begrijpen waarom een getraind systeem tot bepaalde output komt: hoewel deze output nog steeds het resultaat is van toepassing van regels, ontbreekt een duidelijke beschrijving van deze regels en een toelichting op hun relevantie. Ook kan de hoeveelheid en de onderlinge verwevenheid van deze regels het lastig maken om de werking van een systeem te doorgronden.¹⁰ Tot slot kunnen stochastische elementen ingebouwd worden, waardoor de output niet steeds dezelfde is bij dezelfde input (non-determinisme). Noemenswaardig zijn grote taalmodellen (large language models, LLMs); deze zijn vanwege hun enorme omvang, complexiteit en non-determinisme in het bijzonder moeilijk inzichtelijk te maken.¹¹

8 HR 19 juni 2020, ECLI:NL:HR:2020:1082, NJ 2020/15 m.nt. S.D. Lindenbergh (*Miragelplombe*); HR 19 juni 2020, ECLI:NL:HR:2020:1090, NJ 2020/16 m.nt. S.D. Lindenbergh (*PIP*).

9 De 'tenzij'-formule uit art. 6:77 BW komt, hoewel van fundamenteel belang, (nog) niet aan bod.

10 Zie T.F.E. Tjong Tjin Tai, 'Liability for artificial intelligence', in: V. Mak, T.F.E. Tjong Tjin Tai & A. Berlee (red.), *Research Handbook in Data Science and Law*, Cheltenham: Edward Elgar 2024, par. 2 ('What is AI?'); J.A. McDermid, Y. Jia & I. Habli, 'AI for Lawyers. A Gentle Introduction', in: E. Lim & P. Morgan (red.), *The Cambridge Handbook of Private Law and Artificial Intelligence*, Cambridge: Cambridge University Press 2024. Denk ook aan bronpluraliteit, online learning (of zelfs online automated machine learning), ensemble learning en reinforcement learning. Vgl. Tjong Tjin Tai 2024, p. 96-97. Soms is 'post hoc explainability' alsnog deels mogelijk; terechte kritiek daarop is dat dit vaak slechts om lokale verklaringen gaat, hoewel dit volgens S. Steel, 'Legal Causation and AI', in: Lim & Morgan 2024, p. 192 weer niet zo erg hoeft te zijn.

11 A. Holzinger, K. Zatloukal & H. Müller, 'Is Human Oversight to AI Systems Still Possible?', *New Biotechnology* (85) 2025, p. 59-62 (m.n. p. 60).

Ook niet-technische factoren kunnen een rol spelen. Zo kan inzage in de werking van het systeem omwille van bedrijfseconomische belangen of praktische redenen voorbehouden zijn aan eigen medewerkers van de fabrikant.¹² Ook kunnen bij de totstandkoming van het systeem veel verschillende partijen betrokken zijn geweest, waardoor het lastig kan zijn om individuele bijdragen van elkaar te onderscheiden ('many hands'-probleem).¹³ Deze niet-technische factoren zijn echter niet uniek voor AI-systemen.¹⁴

Bovenstaande oorzaken voor 'ondoorgroendelijkheid' en onvoorspelbaarheid vertalen zich op verschillende manieren in juridische uitdagingen. Dit raakt vooral de patiënt als doorgaans eisende partij.¹⁵ Een eerste kwestie betreft normschendingen die *voorafgaan aan de AI-output*, zoals het invoeren van verkeerde data (bijvoorbeeld in weerwil van gebruiksinstructies). Het voorstel voor een AI-aansprakelijkheidsrichtlijn (AILD) zag op uitdagingen van deze aard,¹⁶ maar is inmiddels ingetrokken.¹⁷

Een tweede kwestie, die speelt bij de *weging en interpretatie* van AI-output, betreft de uitdaging om te definiëren wat bij dit gebruik van AI-output als normschending heeft te gelden.¹⁸ Deze kwestie wordt slechts zeer beperkt behandeld in recente AI-gerichte EU-regelgeving.¹⁹ De norm voor de hulpverlener is en blijft 'goed hulpverlenerschap' (art. 7:453 BW), net als wanneer er met andere, 'gewone' technologie gewerkt wordt.²⁰ De toepassing ervan – namelijk wat in een concreet geval van hulpverleners verwacht mag worden – is echter wel afhankelijk van contextuele

12 Vgl. Tschider 2024, p. 342-350.

13 B. Bartlett, 'The possibility of AI-induced medical manslaughter: Unexplainable decisions, epistemic vices, and a new dimension of moral luck', *Medical Law International* (23) 2023, afl. 3, p. 241-27; B. Chan, 'Applying a common enterprise theory of liability to clinical AI systems', *American Journal of Law & Medicine* (47) 2021, afl. 4, p. 351-385; H. Smith, G. Birchley & J. Ives, 'Artificial intelligence in clinical decision-making: Rethinking personal moral responsibility', *Bioethics* (38) 2024, afl. 1, p. 78-86.

14 Vgl. H. Nissenbaum, 'Accountability in a Computerized Society', *Science and Engineering Ethics* (2) 1996, p. 25-42 en A. Feder Cooper e.a. 'Accountability in an Algorithmic Society: Relationality, Responsibility, and Robustness in Machine Learning', *ACM FAccT '22*, p. 864-876 (m.n. par. 3.1). Denk ook aan situaties waarvoor productaansprakelijkheid in het leven is geroepen en recente literatuur over aansprakelijkheid bij zorgnetwerken. Deze uitdagingen zijn bij AI soms wel groter, zie V.P. Maroudas, 'Fault-Based Liability for Medical Malpractice in the Age of Artificial Intelligence', *Review of European and Comparative Law* (57) 2024, afl. 2, p. 135-169 (m.n. p. 140-141).

15 Vgl. Smith, Birchley & Ives 2024; A.A. Fainmann, 'The problem with Opaque AI', *The Thinker* (82) 2019, afl. 4, p. 44-55.

16 Voorstel voor een Richtlijn betreffende de aanpassing van de regels inzake niet-contractuele civielrechtelijke aansprakelijkheid aan artificiële intelligentie (2022/0303/COD, COM(2022) 496 final).

17 J. De Bruyne e.a., 'Briefs on Artificial Intelligence and Liability. Some Points of Attention and Future Perspectives', *CiTiP Working Paper 2025*, p. 11-19.

18 Vgl. Tjong Tjin Tai 2024, par. 5.

19 Vgl. T.J. de Graaf, 'Buitencontractuele aansprakelijkheid van producenten van AI-systemen', *NTBR* 2023/42, par. 5.

20 Vgl. Van Kolfshoeten & Wijne 2024; over kwaliteitsstandaarden C.B.M.C. Zegveld & J.F.L. van de Sande, 'Wie bepaalt wat goede zorg is?', *TvGR* 2024, p. 401-412; over het risico van 'deskilling' M.C. Ploem & M.A.R. Bak, 'Naar een technologiebestendige WGBO', in: *Een toekomstbestendige WGBO. Preadvies Vereniging voor Gezondheidsrecht*, Den Haag: Boom 2025, par. 3.3.3.

factoren, waaronder welke invloed AI-technologie heeft op het medische beslissingsproces en de ruimte die een hulpverlener daarbij praktisch heeft of ervaart.

2.2 Zorgen over een 'liability gap'

Zo kan voor de patiënt een probleem ontstaan. Hulpverleners kunnen niet geacht worden alle ins en outs van een AI-systeem te kennen en altijd alle output te dubbelchecken. Dit zou volgens sommigen een 'aansprakelijkheidsleemte' ('liability gap') opleveren. Indien de keuze voor een bepaald AI-systeem niet per se al een normschending oplevert,²¹ zou de hulpverlener enige mate van vertrouwen moeten mogen hebben in de AI-output,²² en als van een hulpverlener niet verwacht wordt dat hij alle output dubbelcheckt kan dat ook geen onderdeel zijn van diens zorgvuldigheidsnorm.²³ Daardoor ontstaat een ruimte waarin een AI-systeem invloed kan hebben op het beslissingsproces, maar deze invloed niet onder controle van de hulpverlener staat. In die zin opereert het systeem in zekere mate 'autonoom' en heeft het in termen van de principaal-agenttheorie een eigen 'agency' (handelingsbevoegdheid). Auteurs als Duffourc waarschuwen dat voor deze handlingsruimte noch het systeem zelf, noch een ander verantwoordelijk kan worden gehouden. Dit zou een leemte in het aansprakelijkheidssysteem kunnen creëren ('liability gap').²⁴

Let wel: het gaat hier niet om een juridisch begrip van agency zoals bij vertegenwoordiging, lastgeving of agentuur. Het is veeleer een gevolg van de feitelijke capaciteit van het systeem om bepaalde taken geautomatiseerd uit te voeren, waarbij het zelf niet aansprakelijk kan zijn bij gebrek aan rechtspersoonlijkheid en de hulpverlener niet aansprakelijk is omdat hij op de output mocht vertrouwen. Een productaansprakelijkheidsclaim zou, althans in de VS,²⁵ afstuiten op de 'learned

21 Vgl. Maroudas 2024, par. 4.1-4.3.

22 '[T]he choice to use AI in the first place puts the doctor in the position of believing it or not almost as an article of faith' (Selbst 2020, p. 1339); H. Smith & K. Fotheringham, 'Artificial intelligence in clinical decision-making: rethinking liability', *Medical Law International* (20) 2020, afl. 2, p. 133-136; G. Contissa, 'Legal and governance perspectives on meaningful human control: the case of automated mobility', in: G. Mecacci e.a. (red.), *Research Handbook on Meaningful Human Control of Artificial Intelligence Systems*, Cheltenham: Edward Elgar 2024, p. 87-88 en 98-99.

23 Vgl. Smith, Birchly & Ives 2022, p. 80; B. Bartlett, 'Clinical negligence in an Age of Machine Learning: res ipsa loquitur to the Rescue?', *Journal of European Tort Law* (15) 2024, afl. 3, p. 302-305; althans niet zonder deze vergaand te objectiveren; onmogelijk is dat echter niet. Vgl. de annotatie van Lindenbergh onder de PIP- en Miragelplombe-arresten in NJ 2021/16, par. 8 en literatuur over art. 7:658 BW.

24 M.N. Duffourc, 'Malpractice by the Autonomous AI Physician', *University of Illinois Journal of Law, Technology, & Policy* (23) 2023, afl. 1, par. I, onder E; N. Diakopoulos, 'Transparency', in: M.D. Dubber, F. Pasquale en S. Das (red.), *The Oxford Handbook of Ethics of AI*, Oxford: Oxford University Press 2020, p. 209 en ook daarin V. Dignum, 'Responsibility and Artificial Intelligence', p. 219; Bartlett 2024, p. 297; G. Contissa, 'Legal and governance perspectives on meaningful human control: the case of automated mobility', in: G. Mecacci e.a. (red.), *Research Handbook on Meaningful Human Control of Artificial Intelligence*, Cheltenham: Edward Elgar 2024, par. 6.2.2.

25 Ik meen dat dit in de EU genuanceerder ligt, zie J.H. van Staalduinen, 'European Product Liability for AI-based Clinical Decision Support Systems', in: K. Prifti e.a. (red.), *Digital Governance*, Cham: Springer 2024.

intermediary'-doctrine en schuld-gerelateerde elementen in dit regime die een drempel opwerpen bij onzekere AI-risico's.²⁶

Men zou deze aansprakelijkheidsleemte kunnen wegedeneren via de stelling dat het bij juridische verantwoordelijkheid, in het bijzonder civiele aansprakelijkheid, niet zozeer gaat om het toerekenen van handelingen, maar om verdeling van risico's en dat deze risico's bij wijze van uitgangspunt en bij afwezigheid van een aansprakelijkheidsgrondslag nu eenmaal binnen de risicosfeer vallen van degene die feitelijk met de schade geconfronteerd wordt (de patiënt).²⁷ Toch kan men ook dan de vraag stellen of het wel rechtvaardig is om onzekerheden rond geavanceerde AI-systemen altijd voor rekening van de patiënt te laten komen, met name als de voorspelling zou uitkomen dat dergelijke systemen in de toekomst een grotere rol gaan spelen in het medische beslissingsproces. In dit artikel wil ik geen standpunt innemen over hoe wenselijk het is om deze risico's voor rekening van de patiënt te laten komen. Ik ga hieronder echter wel in op de daaraan voorafgaande vraag hoe het Nederlandse recht deze risicosferen nader afgrenst en, dus, hoe het met de door onder meer Duffourc gesignaleerde 'liability gap' omgaat.

Een andere, meer medisch-inhoudelijke juridische kwestie die voortvloeit uit deze 'liability gap' wordt naar voren gebracht door Price, Gerke en Cohen.²⁸ Zij waarschuwen voor prikkels in het aansprakelijkheidsrecht die hulpverleners ertoe aanzetten AI slechts te gebruiken ter bevestiging van wat ze toch al zouden hebben gedaan ('confirmatory advice'). Het medisch aansprakelijkheidsrecht zet volgens hen aan tot conservatief gedrag, nu het opvolgen van het AI-advies 'standaard-medicijn toedienen' normaliter niet kan leiden tot aansprakelijkheid – want: óf geen schade, óf wel schade maar geen normschending – terwijl het opvolgen van het advies 'niet-standaardmedicijn' eerder tot aansprakelijkheid leidt – want: óf geen schade, óf wel schade en normschending.²⁹ Als buiten de grenzen van de zorgvuldigheid geen aansprakelijkheid voor de hulpverlener kan ontstaan, bevat het aansprakelijkheidsrecht aldus een intrinsieke prikkel om de bestaande 'standard of care' te handhaven. De toegevoegde waarde van AI-systemen zit hem volgens deze auteurs vooral in het geven van 'non-standard' adviezen en als hulpverleners ontmoedigd worden deze op te volgen, zouden de voordelen van AI onvoldoende worden benut.

Hieronder wordt geanalyseerd hoe deze twee zorgen zich verhouden tot de Nederlandse contractuele verplichtingen van de hulpverlener, om ten slotte voorzichtig

26 Duffourc 2023, par. I, onder E.

27 Vgl. T. Hartlief, 'Ieder draagt zijn eigen schade', *RM Themis* 2024/1, p. 1-7.

28 W.N. Price II, S. Gerke & I.G. Cohen, 'Liability for use of artificial intelligence in medicine', in: B. Solaiman & I.G. Cohen (red.), *Research Handbook on Health, AI and the Law*, Cheltenham: Edward Elgar 2024, p. 150-166, par. 2. Vgl. ook M.N. Duffourc & S. Gerke, 'The proposed EU Directives for AI liability leave worrying gaps likely to impact medical AI', *NPJ Digital Medicine* (6) 2023, afl. 1.

29 Ze lijken de term 'standaard' te gebruiken in de zin van de/een abstracte professionele standaard ('standard of care') en niet een concrete kwaliteitsstandaard/protocol/richtlijn, maar dat wordt (mij) niet helemaal duidelijk. De laatstgenoemde betekenis zou namelijk wel weer beter passen in hun redenering; vgl. de kritiek in par. 5 hieronder.

te concluderen dat de risico's die Duffourc en Price e.a. signaleren bij ons minder groot zouden kunnen zijn.

3. Zorgvuldigheid en ongeschiktheid vanuit juridisch-dogmatisch perspectief

3.1 *Zorgvuldigheid en ongeschiktheid in wet en jurisprudentie*

De verhouding tussen patiënt en hulpverlener is (in de reguliere zorg) naar Nederlands recht van contractuele aard.³⁰ Het betreft een geneeskundige behandelingsovereenkomst, die als bijzondere overeenkomst een eigen regeling in het Burgerlijk Wetboek (BW) kent (afd. 7.7.5 BW). Vanwege de gelaagde structuur van het BW zijn daarbij ook de algemenere regels uit met name Boek 3 en Boek 6 BW relevant. De verplichtingen die uit hoofde van afdeling 7.7.5 BW op hulpverlener en patiënt rusten, zijn civielrechtelijk te kwalificeren als verbintenissen.³¹ Ze kunnen ex artikel 6:248 lid 1 BW worden aangevuld. Bij een tekortkoming in de nakoming van deze verbintenissen heeft de schuldeiser ex artikel 6:74 BW recht op schadevergoeding, tenzij de schuldenaar zich kan beroepen op overmacht (art. 6:75 BW). Tekortkomingen die ontstaan doordat met een ongeschikte hulpzaak wordt gewerkt, worden toegerekend tenzij dat gelet op de inhoud en strekking van de geneeskundige behandelingsovereenkomst, de verkeersopvattingen of overige omstandigheden onredelijk zou zijn (art. 6:77 BW).

Over deze bepalingen heeft de Hoge Raad in 2020 twee veelbesproken arresten gewezen. In het *Miragelplombe*- en het *PIP*-arrest werd allereerst benadrukt dat later opgekomen inzichten niet meegerekend mogen worden bij het bepalen of er op het moment van behandeling sprake was van een tekortkoming.³² Het bleef echter onduidelijk welk niveau van maturiteit geëist wordt van nieuwe inzichten om wel te mogen meewegen.³³ Daarnaast bleef onduidelijk welke rol 'goed hulpverlener-schap' precies heeft bij het bepalen van ongeschiktheid.³⁴ Met name de rechtsoverwegingen in het *PIP*-arrest riepen hierover vragen op, omdat van een hulpverlener praktisch gezien toch niet kon worden verwacht dat deze de chemische samenstelling van borstimplantatenvulling controleerde.³⁵ Hoewel dat niveau van inspanning niet werd verwacht, kon ondeugdelijke implantaatvulling toch tot een tekortkoming leiden.

30 Zie over buitencontractuele aansprakelijkheid (art. 6:162 BW) en bewusteloze patiënten L.A.B.M. Wijntjens e.a., 'Hoever reikt de persoonlijke aansprakelijkheid?', *NTBR* 2023/37, par. 4.2-4.3; vgl. HR 18 september 2015, ECLI:NL:HR:2015:2745, NJ 2016/66, m.nt. Van Schilfgarde (*Breeweg/Wijkamp*).

31 Ook de relatie als zodanig kan echter als 'verbintenis' worden gekwalificeerd. Zie J.M. Smits, *Bronnen van Verbintenissen* (Mon. Nieuw BW, deel A2), Deventer: Kluwer 2003, hoofdstuk II.

32 HR 19 juni 2020, ECLI:NL:HR:2020:1082, NJ 2020/15 m.nt. S.D. Lindenbergh (*Miragelplombe*); HR 19 juni 2020, ECLI:NL:HR:2020:1090, NJ 2021/16 m.nt. S.D. Lindenbergh (*PIP*).

33 Annotatie G.M. Veldt bij de arresten *PIP* en *Miragelplombe*, *TvC* 2020, 6, p. 353; P.W.J. Verbruggen, 'De maatstaf van state of the art bij aansprakelijkheid voor ongeschikte hulpzaken', *NTBR* 2022/12.

34 R.P. Wijne, 'Langverwachte (on)duidelijkheid over de (on)redelijkheid van aansprakelijkheid van de hulpverlener voor het gebruik van medische hulpzaken', *MvV* 2020, p. 354-363 (m.n. p. 359).

35 Annotatie G.M. Veldt bij de arresten *PIP* en *Miragelplombe*, *TvC* 2020, 6, p. 353.

3.2 Niet gelijk...

Mijns inziens ligt de oplossing niet in een strengere – dat wil zeggen: meer eisend dan praktisch te verwachten – interpretatie van de zorgvuldigheidsplicht, waarop Lindenbergh wel lijkt te hinten.³⁶ Een eenvoudiger verklaring is dat het hier gaat om los te beschouwen verplichtingen. Bij de algemene verbintenis om de zorgvuldigheid van een goed hulpverlener te betrachten, gaat het om een plicht met een inspanningskarakter. Waar ‘de verbintenis in het algemeen [vergt dat; JHS] door de hulpverlener gebruikte zaken (...) voor de toepassing geschikt zijn’ gaat het echter om een verbintenis met een resultaatkarakter waarbij de geleverde inspanning dan wel normatief verwachte inspanning irrelevant zijn.³⁷

Dit valt dogmatisch gezien langs verschillende routes te onderbouwen. Ten eerste zou men, naar recent Belgisch voorbeeld,³⁸ simpelweg kunnen spreken van een aparte (‘veiligheids’-)resultaatsverbintenis binnen de anderszins door inspanningsverbintenissen gekenmerkte behandelingsovereenkomst.³⁹ Ten tweede zou men ‘zorgvuldigheid’ (verwachte niveau van inspanning) los kunnen zien van ‘zorg’ (de toegezegde activiteit, inhoudende ontferming over bepaalde belangen van de schuldeiser).⁴⁰ Het gebruik van een ongeschikte hulpzaak is niet altijd te voorkomen. Dit is alsdan geen tekortschieten in de zorgvuldigheid, maar wel een aantasting van de te leveren (goede) zorg. De hulpzaak wordt niet gebruikt om zorgvuldigheid te betrachten (de maatstaf), maar om te zorgen (de activiteit) en kan dus los van de zorgvuldigheid worden beoordeeld. Op die manier hoeven kwestieuze eigenschappen van hulpzaken, zoals een verkeerde implantaatvulling, niet voor de individuele hulpverlener kenbaar te zijn om toch van een tekortkoming te kunnen spreken.⁴¹

3.3 ...maar wel samenhangend

De geschiktheid van de hulpzaak staat zodoende los van zorgvuldigheid. Toch is deze geschiktheid niet onafhankelijk van de gevraagde activiteit (‘zorg’): zo een hulpzaak ongeschikt is, is hij dat *omdat* de kwestieuze eigenschap deze activiteit,

36 S.D. Lindenbergh in zijn noot in *NJ* 2021/16, par. 8.

37 R.o. 2.8.2 van het *PIP*-arrest.

38 J. van Vaerenbergh, ‘Voortaan Dokter Bibber? Hof van Cassatie bevestigt voor het eerst de veiligheidsverplichting voor medische hulpzaken als resultaatsverbintenis’, *RW* (25) 2024, p. 984-990; S. Stijns, *Leerboek verbintenissenrecht. Boek 1*, Brugge: Die Keure 2025, nr. 201.

39 In België is de wettelijke grondslag daarvoor de evenknie van ons art. 6:248 lid 1 BW, zie Van Vaerenbergh 2024.

40 T.F.E. Tjong Tjin Tai, *Zorgplichten en zorgethiek* (diss. Amsterdam UvA), Deventer: Kluwer 2006, p. 97-100 en 101-122.

41 M.i. is r.o. 5.6 van Hof Den Bosch 18 augustus 2020, ECLI:NL:GHSHE:2020:2592, *JA* 2020/152 m.nt. R.S.B. Hordijk, dan ook te kort door de bocht, hoewel qua voortgeschrijdende medische inzichten wel goed te volgen gezien het arrest *Miragelplombe*. Vgl. R.P. Wijne, ‘Hoge Raad schept duidelijkheid maar beslist is het nog niet’, *L&S* 2020, afl. 3, p. 6-22 (hierna: Wijne 2020a), par. 4.4. Hetzelfde geldt voor r.o. 3.11 van Hof Den Bosch 5 april 2022, ECLI:NL:GHSHE:2022:1080, *JA* 2022/91 m.nt. C.J. van Weering (vervolgarrest *Miragelplombe*).

het verlenen van goede zorg, belemmert.⁴² Eigenschappen die deze activiteit niet verstoren, kunnen ook geen tekortkoming veroorzaken – of, als men uitgaat van een aparte ‘veiligheidsverbintenis’: geen aantasting van de veiligheid vormen. Om deze reden kan de vraag of een hulpzaak ongeschikt is, worden beantwoord door de vraag te stellen wat de maatmens-hulpverlener zou hebben gedaan als hij de kwestieuze eigenschap van de specifieke⁴³ hulpzaak (bijvoorbeeld de kans op scheuren en lekken bij een borstimplantaat,⁴⁴ het vrijkomen van metaaldeeltjes bij een heupimplantaat,⁴⁵ het afglijdpotentieel van een Falope-ring⁴⁶ of de mate van ‘intermittent regurgitation’ van een hartklep⁴⁷) wél zou hebben gekend. Zou hij de hulpzaak alsnog (kunnen) hebben gebruikt? Zo ja, dan is de hulpzaak dus niet ongeschikt en kan het gebruik van de hulpzaak ondanks deze schadeveroorzakende eigenschap naar heersende medische inzichten als ‘goede zorg’ en dus ‘state of the art’ worden gekwalificeerd.⁴⁸ Hierbij kunnen patiënt-specifieke omstandigheden worden meegewogen: deze afweging kan immers bij de ene patiënt(groep) anders uitvallen dan bij de andere.⁴⁹ Dit is wat het volgens mij betekent om geschiktheid ‘in functie van de inhoud van de verbintenis’ te beoordelen.⁵⁰

Voorwaarde voor deze redenering is wel dat de betwiste eigenschap juridisch gezien ‘bestaat’ op het tijdstip van gebruik (vgl. het arrest *Miragelplombe*). Bij kwestieuze eigenschappen die niet of niet langer onderhevig zijn aan voortschrijdende inzichten moet het belang van deze beperking echter niet worden overdreven. Denk hierbij aan PIP-implantaten,⁵¹ een leeggelopen operatiematras,⁵² ondeugdelijke mondkapjes,⁵³ of een hijskraan met losgelopen bout.⁵⁴ Waar voortschrijdende inzichten wel een rol spelen, zoals bij de *Miragelplombe*, of de rol van biofilms bij fillers,⁵⁵ is de discussie in de literatuur over kennisniveaus relevant.⁵⁶ Bij de beoor-

42 Vgl. ‘in functie van de inhoud van de verbintenis’, Concl. A-G Wissink 21 februari 2020, ECLI:NL:PHR:2020:175 (*Miragelplombe*), NJ 2021/15, punt 9.2. Dit sluit ook aan bij het arrest *PIP*, r.o. 2.8.2; vgl. de wethistorische context van art. 6:77 BW zoals beschreven in Hiemstra 2018, par. 4.2-4.3.

43 Hiemstra 2018, par. 4.5.3.3.

44 Arrest *PIP*, r.o. 2.8.4.

45 Rb. Oost-Brabant 17 augustus 2022, ECLI:NL:RBOBR:2022:3395; Rb. Rotterdam 28 juni 2023, ECLI:NL:RBROT:2023:5214, JA 2023/110 m.nt. Klein Gunnewiek.

46 Zie Wijne 2020a, p. 18-19 over HR 1 december 2000, ECLI:NL:HR:2000:AA8724, VR 2001/92, r.o. 3.6.

47 Hof Den Haag 18 september 2012, ECLI:NL:GHSGR:2012:BX8413 (*Advantage*), JA 2012/212 m.nt. J.P.M. Simons, r.o. 8.

48 Arrest *PIP*, r.o. 2.8.2; Hiemstra 2018, par. 4.6.4.5.2.

49 Vgl. de MoM-heupimplantaten in Rb. Oost-Brabant 17 augustus 2022, ECLI:NL:RBOBR:2022:3395, r.o. 5.33 en 5.87-5.92.

50 Wijne 2020a, p. 14 verwijzend naar Concl. A-G Wissink 21 februari 2020, ECLI:NL:PHR:2020:175 (*Miragelplombe*), NJ 2021/15, punt 9.2.

51 Zie R.P. Wijne in A. Kolder (red.), *Handboek Personenschade*, par. 1040.4.2.

52 Rb. Amsterdam 23 december 2021, ECLI:NL:RBAMS:2021:7281.

53 R.P. Wijne, ‘Aansprakelijkheid van de hulpverlener voor COVID-19-gerelateerde schade’, *L&S* 2020/4, par. 4.

54 Rb. Overijssel 2 december 2021, ECLI:NL:ROVE:2021:4915.

55 Hof Den Bosch 18 augustus 2020, ECLI:NL:GHSHE:2020:2592, JA 2020/152 m.nt. R.S.B. Hordijk.

56 Zie o.m. W.H. van Boom, ‘Gebrekkige hulpzaak bij medische behandeling’, *AA* 2021/3, p. 274 en 276-277; Veldt 2020, p. 352-353; Wijne 2020a, p. 14-15.

deling van de geschiktheid van een concreet AI-systeem is het dus van belang om vast te stellen of de betwiste eigenschap hieronder valt. Als en naarmate er door een AI-systeem zelf nieuwe medisch-inhoudelijke verbanden worden ontdekt en toegepast, zou hier eerder sprake van kunnen zijn dan wanneer er 'slechts' met bestaande kennis wordt gewerkt.

4. Zorgvuldigheid en ongeschiktheid vanuit socio-technisch perspectief

4.1 Inleiding

Met het onderscheiden van zorgvuldigheid en het gebruik van geschikte hulpzaken als twee losstaande verplichtingen, komt de vraag op hoe dit onderscheid in de praktijk gemaakt moet worden. De twee verplichtingen kennen immers een andere maatstaf voor het aannemen van een tekortkoming en een ander toerekeningskader (denk met name aan de 'tenzij'-formule in artikel 6:77 BW). Daartoe bespreek ik hieronder twee inzichten uit socio-technisch onderzoek (*sociotechnical systems theory*, STS) naar AI-gebruik:⁵⁷ de 'Level of Automation Taxonomy' (par. 4.2) en de 'role taxonomy' (par. 4.3). De rechtvaardiging hiervoor ligt in de veronderstelling dat de zorgvuldigheidsplicht de praktisch beschikbare handelingsruimte volgt. Ik vat in paragraaf 4.4 samen hoe dit ons kan helpen bij de afbakening.

4.2 Level of Automation Taxonomy

De 'Level of Automation Taxonomy' (LOAT) is gebaseerd op het veelgebruikte glijdendeschaalmodel dat is weergegeven in figuur 1-a.⁵⁸ Met dit model kan worden aangegeven in hoeverre een softwaresysteem 'automatisch' werkt en zich dus zou kunnen onttrekken aan de zorgvuldigheid van de hulpverlener. Systemen kunnen ook onderling vergeleken worden, bijvoorbeeld door het spectrum te schalen naar het betreffende model. In figuur 1-b en 1-c is dit weergegeven voor twee systemen: DECIDE, een risicomodel dat de kans op hart- en vaatziekten inschat en dat centraal staat in het onderzoeksproject waarbinnen dit artikel is geschreven,⁵⁹ en IDx-DR (tegenwoordig: LumineticsCore), dat op basis van een retinascan inschat of de patiënt mogelijk lijdt aan diabetische retinopathie.⁶⁰

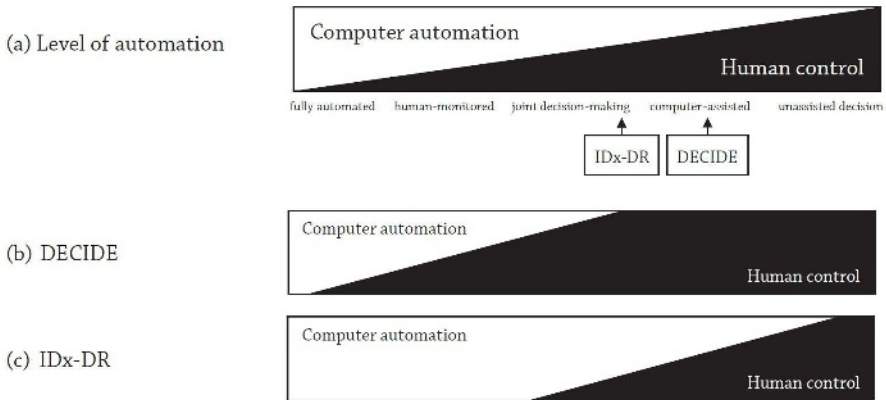
57 Zie over dit onderzoeksveld C.B. Graber, 'How the Law Learns in the Digital Society', *Law, Technology and Humans* (3) 2021, afl. 2, p. 12-27 (m.n. par. 3.1-3.3); H. Kydén, 'Sociology of digital law and artificial intelligence', in: J. Přibáň (red.), *Research Handbook on the Sociology of Law*, Cheltenham: Edward Elgar 2020; A. Beckers & G. Teubner, *Digitale Aktanten, Hybride, Schwärme*, Berlijn: Suhrkamp 2024, par. III.1. Vgl. ook T. Bouwman, 'Empirisch-juridisch onderzoek en de sprong van feit naar norm', *NJB* 2020/815.

58 Zie bijv. de *Leidraad kwaliteit AI in de zorg* (versie 1.1, 2023), p. 50.

59 H.J.A. van Os e.a., 'Cardiovascular Risk Prediction in Men and Women Aged Under 50 Years Using Routine Care Data', *Journal of the American Heart Association* (12), 2023, afl. 7; Project 'DECIDE-VerA', ZonMW nr. 08540122120004. DECIDE staat voor 'Clinical DECision support system for carDiovascular risk managEment in primary care'.

60 Zie Richtlijn Diabetische retinopathie, par. 2.3 'Geautomatiseerde screening middels AI' (versie 20 januari 2023).

Figuur 1 Het spectrum van het algemene 'level of automation'-model (a) vertaald naar twee specifieke AI-toepassingen DECIDE (b) en IDx-DR (c)



In de LOAT-tabel van Lagioia & Contissa is deze glijdende schaal vertaald naar een tabel met afgebakende taken, oplopend in niveau van automatisering en onderverdeeld naar vier fasen van het medisch besluitvormings- en behandelproces: (A) informatievergaring; (B) informatieanalyse; (C) besluitvorming en behandelkeuze; en (D) behandeling (zie tabel 1 hieronder).⁶¹ Dit kan ook inzichten geven voor het aansprakelijkheidsrecht, omdat hiermee een onderscheid gemaakt kan worden tussen schadeoorzaken die verband houden met het gedeelte van het proces waarover de hulpverlener minder tot geen controle heeft en het gedeelte waarover hij dat wel heeft.⁶²

Vergelijk dit met de situatie die in paragraaf 1 werd geschetst: de hulpverlener kan geen onzorgvuldigheid worden verweten als de schadeoorzaak te maken heeft met een aspect van het systeem waar hij geen controle over heeft of waar hij dat wel zou kunnen hebben, maar dit onredelijk veel moeite of tijd kost. Er is dan sprake van een (sub)proces in het systeem dat op volledig automatische wijze en zonder 'human in the loop' werkt. Daarvan kan al sprake zijn bij simpele databasesoftware die patiëntgegevens ophaalt en weergeeft: dit is niet te dubbelchecken door de hulpverlener, althans dat wordt niet van hem verwacht. Bij AI-systemen is het denkbaar dat meer van dit soort (sub)processen zich onttrekken aan een inhoudelijke

61 F. Lagioia & G. Contissa, 'The strange case of Dr. Watson: liability implications of AI evidence-based decision support systems in health care', *European Journal of Legal Studies* (12) 2020, afl. 2, p. 245-289; verder ontwikkeld in Contissa 2024; gebaseerd op L. Save, B. Feuerberg & E. Avia, 'Designing Human-Automation Interaction: A New Level of Automation Taxonomy', in: D. de Waard e.a. (red.), *Human Factors: a view from an integrative perspective. Proceedings HFES Europe Chapter Conference Toulouse 2012*, zie www.hfes-europe.org.

62 Zie nader Contissa 2024, par. 6.2. Hierbij zij opgemerkt dat het gebruik van een AI-systeem ook andere verplichtingen met zich mee kan brengen, zoals het updaten van het systeem (zie aangehaalde literatuur in par. 1).

Mr. drs. J.H. van Staalduin

(dubbel)check van de hulpverlener. De LOAT-tabel helpt om te bepalen welke (sub)processen dit zijn.

Tabel 1 Level of Automation Taxonomy (LOAT)

A. Information acquisition	B. Information analysis	C. Decision and action selection	D. Action implementation
A0. Manual information acquisition	B0. Working-memory based information analysis	C0. Human decision-making	D0. Manual action and control
A1. Artefact supported information acquisition	B1. Artefact supported information analysis	C1. Artefact-supported decision-making	D1. Artefact supported action implementation
A2. Low-level automation support of info acquisition	B2. Low-level automation support of info analysis	C2. Automated decision support	D2. Step-by-step action support
A3. Medium-level automation support of info acquisition	B3. Medium-level automation support of info analysis	C3. Rigid automated decision support	D3. Low-level support of action sequence execution
A4. High-level automation support of info acquisition	B4. High-level automation support of info analysis	C4. Low-level automatic decision making	D4. High-level support of action sequence execution
A5. Full automation support of info acquisition	B5. Full automation support of info analysis	C5. High-level automatic decision making	D5. Low-level automation of action sequence execution
		C1. Artefact-supported decision-making	D6. Medium-level automation of action sequence execution
			D7. High-level automation of action sequence execution
			D8. Full automation of action sequence execution

De tabel kan worden toegepast op een concreet AI-systeem door binnen elke kolom het niveau van automatisering te bepalen. Om hiervan een praktisch voorbeeld te geven, is met de andere onderzoekers binnen het onderzoeksproject DECIDE-VerA overlegd welke score DECIDE zou krijgen. Dit is voor de eerste kolom A4 of A5, afhankelijk van de implementatie, omdat de patiëntgegevens automatisch worden opgehaald. In de tweede kolom scoort het B5, omdat de risicoscores automatisch worden berekend zonder tussenkomst van de hulpverlener. In de derde kolom scoort het C0 of C1, afhankelijk van de inbedding van het systeem, omdat het model zelf geen opties genereert (slechts een risicoscore), maar dit in combinatie met andere elementen in het EPD mogelijk wel als 'artefact-supported' besluitvorming kan gelden. In de vierde kolom scoort het D0, omdat geen hulp wordt geboden bij de al dan niet te ondernemen actie. Deze scores zijn weergegeven in tabel 2a hierna. Op een vergelijkbare manier kan IDx-DR (LumineticsCore) beoordeeld worden (tabel 2b hierna). Opgemerkt zij dat deze score slechts gebaseerd is op openbare informatie; hij dient slechts om te illustreren dat twee systemen op genuanceerde wijze met elkaar vergeleken kunnen worden.

De LOAT-tabel biedt zo, op basis van een taakverdelingsmodel, een manier om onderscheid te maken tussen schadeoorzaken die in de sfeer van de hulpzaak liggen en die welke in de sfeer van zorgvuldigheid liggen. Met de score in de tabel kan, wanneer in een concreet geval schade is ontstaan, bepaald worden welk beoordelingskader (zorgvuldigheid of ongeschikte hulpzaak) relevant is om te beoordelen of er een tekortkoming is gepleegd en hoe vervolgens over toerekening kan worden nagedacht. Als een oorzaak bijvoorbeeld buiten de (redelijkerwijs te verwachten

uitoefening van) controle van de hulpverlener ligt, ligt toepassing van artikel 6:77 BW voor de hand. Bij DECIDE is dit bijvoorbeeld het geval voor oorzaken die gerelateerd zijn aan niveaus A4 (evt. A5), B5 en eventueel C1.

Tabel 2a DECIDE (A4/A5, B5, C0/C1, D0)

A0	B0	C0	D0
A1	B1	C1	D1
A2	B2	C2	D2
A3	B3	C3	D3
A4	B4	C4	D4
A5	B5	C5	D5
		C6	D6
			D7
			D8

Tabel 2b IDx-DR (A5, B5, C4, D0)

A0	B0	C0	D0
A1	B1	C1	D1
A2	B2	C2	D2
A3	B3	C3	D3
A4	B4	C4	D4
A5	B5	C5	D5
		C6	D6
			D7
			D8

4.3 De rol van de 'human in the loop'

Crootof e.a. waarschuwen echter dat het idee dat een werkproces op te delen is in subtaken die hetzij door een AI-systeem, hetzij door een hulpverlener kunnen worden volbracht te reductionistisch is. Een strikte afbakening in subtaken zou vol-

gens hen juist kunnen leiden tot extra gevaar.⁶³ Zij geven als voorbeeld het geval waarin een zelfrijdende auto in complexe wegsituaties de controle teruggeeft aan de bestuurder. Zelfs indien het systeem deze situaties accuraat kan detecteren, schrijven zij, leidt het opdelen in disjuncte subtaken hier tot (extra) risico: als de bestuurder tijdens het automatisch rijden niet aan het opletten is – toch juist het voordeel van zelfrijdende auto's – kan hij ook niet adequaat reageren op plotselinge gevaarlijke situaties.⁶⁴ Een 'taakgericht' model zoals gehanteerd door Lagioia en Contissa zou miskennen dat (sub)taken in een werkelijke toepassings situatie met elkaar zijn verweven.⁶⁵

Zij stellen in plaats daarvan de interactie tussen het AI-systeem en de hulpverlener centraal en analyseren verschillende rollen die een gebruiker hier kan hebben: (1) *corrective*, (2) *resilience*, (3) *justificatory*, (4) *dignitary*, (5) *accountability*, (6) *stand-in*, (7) *friction*, (8) *warm body* en (9) *interface*. Hier zijn met name de *resilience*- en *corrective*-rollen interessant. De eerste houdt in dat de hulpverlener er is om bij bijzondere omstandigheden, zoals storingen, in te grijpen om schade bij patiënten te voorkomen. De tweede houdt in dat de hulpverlener de output van het systeem corrigeert of nuanceert: het kan daarbij gaan om (i) het opmerken van fouten, (ii) het tegenwicht bieden aan bias, of (iii) om 'situational correction', dat wil zeggen het meewegen van contextuele factoren zoals specifieke gevalsomstandigheden (patiënteigenschappen) waar het AI-systeem naar zijn aard minder rekening mee kan houden.⁶⁶

Deze rollen hebben aansprakelijkheidsrechtelijke betekenis in de zin dat uit een bepaalde rol normatieve verwachtingen kunnen voortvloeien, die soms zelfs (iets) verder kunnen gaan dan wat bijvoorbeeld in de beroepsgroep gebruikelijk is. Tegen deze achtergrond kan Lindenberghs interpretatie van het PIP-arrest (zie par. 3.2) worden begrepen. Normatieve verwachtingen mogen ook weer niet te ver afwijken van wat concreet aan inspanning verwacht mag worden, maar ze kunnen wel nuance geven aan het taakgerichte afbakeningsmodel van Lagioia en Contissa.

63 R. Crootof, M.E. Kaminski & W.N. Price II, 'Humans in the loop', *Vanderbilt Law Review* (76) 2023, afl. 2, p. 429-510.

64 Vgl. Hof Arnhem-Leeuwarden 13 december 2022, ECLI:NL:GHARL:2022:10711, JA 2023/29 m.nt. R.S.M. Bosch (automobilist nog steeds bestuurder na aantrekken handrem door inzittende).

65 Zij noemen dit de 'MABA-MABA'-trap: het is in hun ogen niet zinvol om een taak op te delen in gedeeltes waarin de mens beter is ('Men [sic] Are Better At'), zoals improvisatie, contextualisering en waardeoordelen, en een gedeelte waarin AI beter is ('Machines Are Better At'), zoals repetitief werk.

66 Bij medische AI kunnen ook meespelen de *justificatory role*, het legitimeren van een beslissing en toelichten hoe deze tot stand is gekomen; *dignitary role*, het voorkomen van het objectiveren van mensen door ze aan puur algoritmische besluitvorming te onderwerpen en zo te reduceren tot hun 'data shadow'; *interface role*, het als domeinexpert helpen van leken bij het gebruik van het systeem; en *accountability role*, het dragen van morele en juridische verantwoordelijkheid – soms in negatieve zin beschreven als 'liability sponge' of 'moral crumple zone'. Zie Smith, Birchley & Ives 2022, p. 79. Verder beschrijven zij de *stand-in role* (ter verdediging tegenover de buitenwereld dat de verantwoordelijke organisatie concrete stappen heeft genomen ter waarborging van veiligheid); *friction role* (ter vertraging van automatische processen zodat afwijkingen opgemerkt kunnen worden); en *warm body role* (ter behoud van banen).

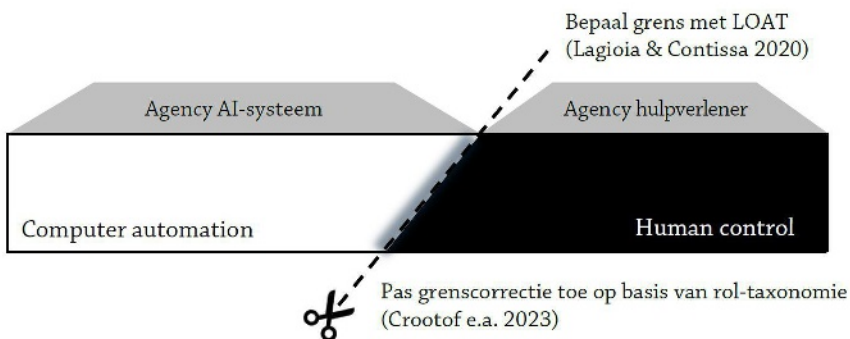
4.4 Synthese: afbakening met taken en rollen

Het *taakgerichte* model van Lagioia en Contissa en het *rolgerichte* model van Crootof e.a. hoeven elkaar volgens mij niet uit te sluiten. De rollen van Crootof e.a. vormen een normatieve grenscorrectie van de taken in het model van Lagioia en Contissa. De verantwoordelijkheid van de gebruiker strekt zich dus iets verder uit dan de grens van zijn taak, hetgeen betekent dat de zorgvuldigheidsnorm dus iets meer kan eisen dan de taak zelf strikt vereist.

Dit gecombineerde model is weergegeven in figuur 2 hierna. Met dit model, gebaseerd op agency van systeem en gebruiker, wordt iets duidelijker hoe hulpzaak en zorgvuldigheid zich bij AI-adviesystemen tot elkaar verhouden. De hulpzaak (links van de grens) wordt beoordeeld volgens de maatstaf die ik beschreef in paragraaf 3.3; de hulpverlener (rechts van de grens) wordt beoordeeld volgens de maatstaf van ‘goed hulpverlenerschap’. Beide zijden van dit model zijn nodig om een oordeel te vellen over de aansprakelijkheid van de hulpverlener, maar ze kennen wel een eigen systematiek.

Ter illustratie: gaat het bij een behandeling, waarbij het DECIDE-systeem is gebruikt, fout bij de informatieanalyse (ten aanzien van de data waar het systeem over gaat), dan zien we in de LOAT-tabel dat dit om kolom B gaat en dat de hulpverlener dat niet met zorgvuldigheid had kunnen voorkomen. Om deze reden moet hier de beoordelingsmaatstaf voor ongeschikte hulpzaken worden toegepast (zie par. 3.3) en verloopt toerekening via artikel 6:77 BW, waardoor rekening moet worden gehouden met de tenzij-formule. In vergelijking met het systeem IDx-DR wordt het interessant als het mis gaat in de besluitvorming en behandelkeuze (kolom C). Bij DECIDE is hier meer ruimte voor zorgvuldigheid van de hulpverlener dan bij IDx-DR, waardoor bij DECIDE minder snel gesproken kan worden van een tekortkoming die terug te voeren is op het gebruik van een ongeschikte hulpzaak.

Figuur 2 ‘Taaktaxonomie’ en ‘roltaxonomie’ als basis voor de afbakening



5. Zorgen over een ‘liability gap’ opnieuw beschouwd

In paragraaf 2.3 besprak ik de zorgen van Duffourc en Price e.a. De laatsten geven zelf al aan dat hun scenario’s abstraheren van de werkelijke situatie. Niettemin schrijven McCradden en Stedman dat een dergelijk binair model geen recht doet aan medische besluitvorming, nu het daarbij niet gaat om juiste of onjuiste beslissingen, maar om *redelijke* beslissingen.⁶⁷

Ook in recente Nederlandse literatuur wordt benadrukt dat het werk van hulpverleners meer omvat dan het stellen van de juiste diagnose (‘een arts is geen diagnosemachine’): dat bestaat namelijk ook uit het meewegen van patiënt-specifieke belangen en voorkeuren en overleg daarover met de patiënt (‘shared decision-making’).⁶⁸ Het argumentatiemodel van Price lijkt aan te nemen dat altijd een ‘juiste’ behandeling – een soort ‘ground truth’ – vast te stellen is. Als er een standaardbehandeling is, zal die doorgaans gevolgd moeten worden en speelt dat een grote rol in de beoordeling achteraf,⁶⁹ maar de taak van de hulpverlener komt uiteindelijk neer op het interpreteren van de beschikbare informatie, waaronder de output van het AI-systeem,⁷⁰ in het licht van het complex aan relevante factoren. *Dat* vormt dan ook het onderwerp van toetsing in het kader van goed hulpverlenerschap: er wordt niet getoetst op de uitkomst van de beslissing, maar of deze beslissing, gegeven de beschikbare informatie, ook door een redelijk bekwaam en redelijk handelend vakgenoot zou kunnen zijn genomen.⁷¹

Toch betekent dit op zichzelf nog niet dat de inzichten van Price e.a. helemaal terzijde moeten worden geschoven. Samen met Duffourc illustreren ze welke prikkels en leemtes een zuivere inspanningsplicht in het leven zou kunnen roepen. Het Nederlandse systeem is echter niet zuiver op inspanningsplichten gebaseerd, waardoor de risico’s die zij signaleren, afhankelijk van hoe de tenzij-formule van artikel 6:77 BW toegepast wordt, iets minder groot zijn. Een hulpverlener kan hier te lande immers ondanks zorgvuldig handelen een tekortkoming plegen. Schade, zoals gezondheidsverlies die op die wijze ontstaat, valt in principe (tenzij dat, in de zin van artikel 6:77 BW, onredelijk wordt geacht) in de risicosfeer van de hulpverlener. Dit geldt gezien het stelsel van artikel 6:74-6:77 BW zowel voor schade die

67 M.D. McCradden & I. Stedman, ‘Explaining decisions without explainability? Artificial intelligence and medicolegal accountability’, *Future Healthcare Journal* (11) 2024, afl. 3, p. 2: ‘[d]ecisions can be reasonable, even when they lead to a bad outcome’.

68 K.R. Jongasma, W.W. van Solinge & S. Haitjema, ‘Acht misvattingen over AI in de zorg’, *NTVG* (167) 2023, (die hiervoor verwijzen naar M. Mittelman, S. Markham & M. Taylor, ‘Patient commentary: Stop hyping artificial intelligence—patients will always need human doctors’, *BMJ* (363) 2018); zie tevens M.C. Ploem, *Goed recht voor zorgtechnologie* (oratie UvA), Den Haag: Boom juridisch 2023, p. 30.

69 Vgl. de ‘protocolarresten’, HR 2 maart 2001, ECLI:NL:HR:2001:AB0377 (*Protocol I*), NJ 2001/649 m.nt. J.B.M. Vranken & F.C.B. van Wijmen; HR 1 april 2005, ECLI:NL:HR:2005:AS6006 (*Protocol II*), NJ 2006/377 m.nt. H.J. Snijders.

70 Wederom onder de aanname dat het gebruik van het systeem per se nog geen tekortkoming oplevert.

71 Zie hierover ook R.W.M. Giard, ‘Oordelen over personenschade veroorzaakt door diagnostische fouten’, *TvVP* 2014, p. 18-23. Hiermee bedoel ik niet te impliceren dat het een terughoudende (marginale) toets zou zijn.

direct door de hulpzaak wordt veroorzaakt, als schade die veroorzaakt wordt doordat een hulpverlener door de ongeschiktheid gestoord wordt in het medische beslissingsproces.

6. Conclusie

In deze bijdrage ben ik ingegaan op de zorgplicht van hulpverleners bij het gebruik van AI-output. Ik begon met de kenmerken van (medische) AI en enkele zorgen die daaruit voortvloeien. De zorg van Duffourc (par. 2.2) hield in dat AI-systemen vanwege hun ondoorgrondelijkheid maar beperkt zijn te controleren, daardoor buiten de zorgplicht vallen en zo eigen 'agency' kunnen krijgen die resulteert in een aansprakelijkheidsleemte.

Ik heb daartoe uiteengezet hoe het Nederlandse stelsel van wanprestatie met risicotoerekening begrepen kan worden. De vereiste zorgvuldigheid en de vereiste geschiktheid van hulpzaken vormen het onderwerp van twee onderscheiden, maar samenhangende verplichtingen. Op deze manier wordt de door Duffourc geschetste leemte gedeeltelijk opgevuld en tegelijk tegemoetgekomen aan de zorg van Price e.a., omdat deze oorzaken alsnog binnen de risicosfeer van de hulpverlener worden gebracht. In hoeverre dit in werkelijkheid ook de uitkomst is, hangt wel af van hoe 'ongeschiktheid' en de reikwijdte van de tenzij-formule in concreto worden toegepast.⁷² De aangehaalde zorgen uit de medische AI-literatuur kunnen daarbij betrokken worden.

Bij de beoordeling van de aansprakelijkheid voor schade waarbij onjuiste AI-output een rol heeft gespeeld, rijst de praktische vraag of deze aan de onzorgvuldige handelwijze van een hulpverlener kan worden toegeschreven dan wel voortvloeit uit de ongeschiktheid van het AI-systeem als hulpzaak. Deze afbakeningsvraag is relevant, omdat (de geschiktheid van) een hulpzaak anders wordt beoordeeld dan (het gedrag van) een hulpverlener. Daartoe heb ik beschreven hoe bijdragen uit de socio-technische literatuur inzicht kunnen geven in de grens van zorgvuldig handelen; aan de andere kant van deze grens geldt het regime voor hulpzaken. Van belang is eerst vast te stellen welke taken en rollen in een concreet geval aan de hulpverlener kunnen worden toegeschreven, hetgeen praktisch te bepalen is met het aan STS-onderzoek ontleende denkkader dat ik besprak in paragraaf 4. Vervolgens kan de juiste beoordelingsmaatstaf worden toegepast. Aldus beschouwd is het Nederlandse aansprakelijkheidsrecht *future proof* en flexibel genoeg om veranderingen in de rollen van arts en AI het hoofd te bieden.

72 Zie hierover o.m. P.W.J. Verbruggen, 'Aansprakelijkheid voor ongeschikte medische hulpzaken volgens de Hoge Raad', *WPNR* 2020/7303; Wijne 2020c; Van Boom 2021; K.A. Jelsma & J. Booij, 'Gebruik van (ongeschikte) hulpzaken: What's next?', *AV&S* 2021/17; T. Hartlief, 'Aansprakelijkheid & Verzekering anno 2023', *AV&S* 2023/32, par. 11.