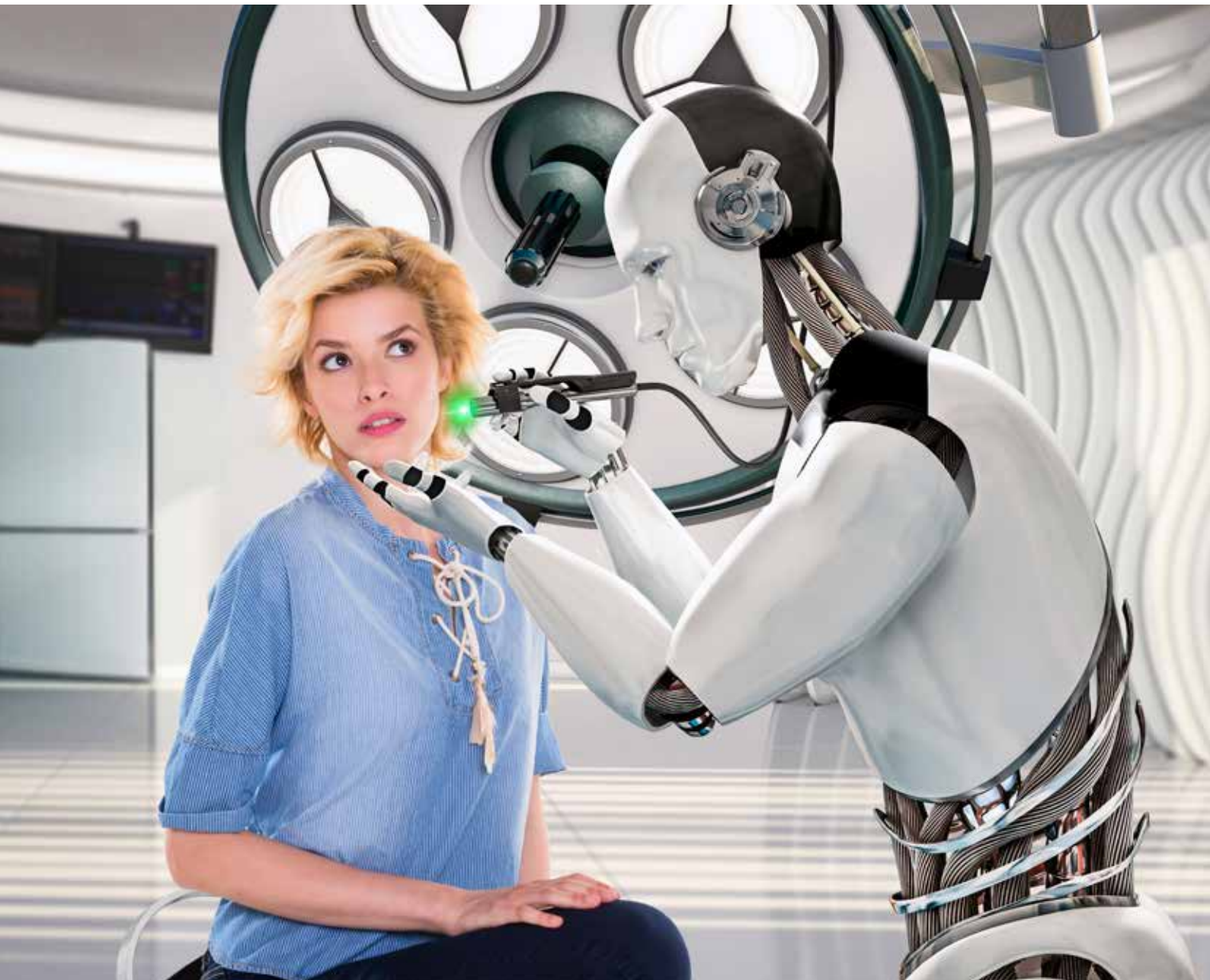


De computer als dokter



Kunstmatige intelligentie lijkt ineens overal op te duiken. Wat is het eigenlijk, wat kun je ermee in de geneeskunde, en wat niet? De mogelijkheden en valkuilen in vier vragen.

1

Je hoort voortdurend over AI, algoritmen, neurale netwerken en *machine learning*. Wat zijn de verschillen en overeenkomsten tussen al die termen?

Kunstmatige intelligentie – of AI (*artificial intelligence*) – is de overkoepelende term voor een variëteit aan technieken waarmee intelligentie wordt gesimuleerd op computers. Grosso modo zijn er twee soorten: regelgebaseerde en zelflerende technieken.

In het eerste geval spreek je meestal van een *expertsysteem*. Meteen maar twee voorbeelden: NHGDoc, dat de huisarts op basis van de NHG-Standaarden ondersteunt bij het geven van advies aan de patiënt, en Oncoguide, dat via beslisbomen patiëntspecifieke behandeladviezen genereert op basis van (ziekte-) gegevens van de patiënt uit het elektronisch patiëntendossier en Nederlandse oncologische richtlijnen. In beide en gelijksoortige gevallen volgt het systeem een algoritme van het type ‘als dit het geval is, doe dan dit of kies dan dat’. Het is – zou je kunnen zeggen – een geprogrammeerde redeneermechanisme die wordt gebruikt om conclusies te trekken in een specifiek klinisch geval. Een voorwaarde voor het ontwikkelen van dit soort systemen is uiteraard dat er consensus bestaat over bepaalde kennis en medische principes, zodat die in regels vastgelegd kunnen worden.

Een veel geavanceerdere vorm van AI is *machine learning*: hier is niets voor geprogrammeerd. Hier gaat het om een systeem dat ‘leert’ op basis van patronen in de data waarmee het wordt gevoed. Op basis daarvan stelt de computer een set van regels of instructies op, die hij vervolgens kan toepassen op nieuwe situaties: een algoritme. De meeste toepassingen

in de zorg vallen onder deze vorm van AI. Er zijn twee hoofdvarianten.

Ten eerste: *supervised learning*, de computer wordt ‘getraind’ met data en feedback, bijvoorbeeld met beelden van huidafwijkingen. Deze data krijgen een label van de desbetreffende huidafwijking. Indien deze gelabelde beelden maar vaak genoeg worden aangeboden aan de com-

Het gaat om een systeem dat ‘leert’ op basis van patronen in data

puter, leert hij dit type huidafwijking herkennen op beelden die nog niet eerder aan de computer zijn aangeboden. Specifiek voor deze manier van leren is dat het juiste antwoord reeds bekend is; de arts heeft de data gelabeld en de computer kan zijn voorspelde uitkomst controleren met de gelabelde uitkomst. Hierdoor leert de computer en stelt zijn eigen regels bij.

Ten tweede: *unsupervised learning*, nu zijn de data niet vooraf gelabeld, geclassificeerd of gecategoriseerd. Deze manier van *machine learning* wordt in de praktijk gebruikt om grote hoeveelheden data te groeperen of om in de data nieuwe patronen te herkennen, die mensen niet

eerder hebben herkend. In feite simuleert dit ook tot op zekere hoogte hoe mensen zelf leren. We hebben immers geleerd om het begrip 'stoel' tot een paar essentiële eigenschappen te abstraheren om zo aan de hand daarvan een nieuw object als 'stoel' te herkennen.

Een speciale vorm van (*un-*)*supervised learning* is *deep learning*; de AI-techniek waarvan het meest wordt verwacht. Ook hier is er niets voorgeprogrammeerd. De werking is eigenlijk afgekeken van de manier waarop onze hersenen functioneren en berust op in de computer gesimuleerde neurale netwerken. Die bestaan uit lagen van virtuele cellen (qua functioneren enigszins vergelijkbaar met hersencellen). Binnen het neurale netwerk wordt informatie via verbindingen van een bepaalde laag naar andere lagen doorgegeven, zoals dat ongeveer ook in ons brein gebeurt. Elke verbinding heeft een instelbaar 'gewicht', wat betekent dat een signaal op volle sterkte of zwakker wordt doorgegeven en of het signaal activerend of juist afremmend uitwerkt. Het neurale netwerk leert een taak, bijvoorbeeld tumoren opsporen op scans, door al die gewichten optimaal in te stellen op basis van een zeer groot aantal aangeboden beelden, met telkens feedback hoe goed het een tumor heeft opgespoord.

2

Hoe komt het dat AI juist nu zo'n stormachtige ontwikkeling lijkt door te maken?

De term artificiële intelligentie duikt voor het eerst op in de jaren vijftig. Het eerste, rudimentaire neurale netwerk dateert zelfs al van 1951 (voor de liefhebber: ontworpen door Marvin Minsky, toen nog student, later een van de grondleggers van AI). Er ging lange tijd een grote belofte van uit, maar pas de laatste decennia kon die belofte worden ingelost. Volgens Pim Haselager, psycholoog, filosoof en AI-deskundige

mogelijk, die noodzakelijk was voor levensechte games. Haselager: 'Die GPU's zijn rekenmonsters. Begin deze eeuw konden neurale netwerken gaan draaien op die GPU's, waardoor opeens de kracht van die benadering zichtbaar werd. Ze werden praktisch bruikbaar.' Een andere, zeer belangrijke ontwikkeling was de komst van het internet: 'Zo kwamen er grote verzamelingen data beschikbaar die nodig zijn om neurale netwerken te voeden.'

'De game-industrie droeg bij aan AI'

aan het Donders Instituut van de Radboud Universiteit, dankzij een paar ontwikkelingen die tezamen de hoge vlucht van AI hebben bevorderd. Ten eerste was er de invloed van de computergamesindustrie die om steeds meer rekenkracht vroeg, en dus om geavanceerdere hardware. De komst van *graphical processing units* (GPU's) maakte de aansturing van de grote hoeveelheid beeldpixels



3

Wat kan de zorg nu al met AI en wat (nog) niet?



Mensen, en dus ook dokters, zijn traag en worden moe. Computers niet: die zijn snel en gaan altijd maar door. Mensen kunnen weliswaar patronen ontdekken in een brei van ongestructureerde data, maar *machine learning* kan met enorm veel grotere databestanden omgaan en daar patronen in vinden. In beide aspecten zit voorlopig de winst van AI-toepassingen in de kliniek. Zo kan een AI-systeem gigantische hoeveelheden röntgenfoto's, MRI's of CT-scans 'bekijken' om tumoren te leren onderscheiden, en in theorie daardoor trefzekerder worden dan radiologen. Precies dankzij de beschikbaarheid van de ingrediënten die we eerder noemden: enorme databases die online beschikbaar zijn, snelle grafische chips in computers, en neurale netwerken. Willekeurig voorbeeld: Nature publiceerde twee jaar geleden een onderzoek waarin een AI-systeem 757 verschillende huidziekten volledig automatisch op basis van visuele informatie diagnosticeerde en in dat opzicht veel beter presteerde dan de 21 dermatologen die aan het onderzoek deelnamen. AI kan ook helpen bij het ontwerpen van medicijnen. Voorbeeld: recent onderzoek laat zien dat de laatste vijftig jaar meer dan 450 geneesmiddelen zijn teruggetrokken van de markt vanwege ernstige, niet voorziene bijwerkingen, vooral door levertoxiciteit. Het metabolisme van de lever is notoir complex, en data over die complexiteit in samenhang zien met data over de toxiciteit van bepaalde moleculen is een kolfje naar de hand van *machine learning*, mits er maar voldoende gegevens beschikbaar zijn. En die zijn er volgens Amerikaanse onderzoekers. Elders in dit nummer nog veel meer voorbeelden van wat AI vermag. Maar AI kent ook – althans nu – grenzen. Zo zijn contextgevoeligheid en 'uitzonderingen op de regel' grote uitdagingen, weet Pim Haselager. 'De AI-successen halen de krant, terwijl de gebieden met stilstand eigenlijk veel interessanter zijn. Mensen kennen zo iets als common sense. Dat is heel moeilijk om in AI-systemen te simuleren. Hetzelfde geldt voor sociale intelligentie, gevoel, beleving – het voortdurende besef van 'er zijn', dat is ook iets wat die systemen

missen. Ik bedoel: een schaakcomputer weet niet wat winnen of verliezen is. Terwijl dat voor ons juist een reden is om te spelen. Dat betekent dat je op sommige plekken AI heel goed kunt inzetten en op andere juist helemaal niet. Daar is verwarring over, ook in de zorg. De robot als gezelschap bijvoorbeeld vind ik een nogal gênante vertoning. Die robot beschikt over een soort standaardrepertoire, maar heeft totaal geen sociale capaciteiten. En die komen er voorlopig ook niet. Je kunt wel proberen hem alle mogelijke sociale interacties te laten berekenen, maar die nemen al gauw exponentieel toe. Hoe snel computers ook zijn, dat lukt ze niet. De bottleneck is dus van psychologische aard: hoe komt het dat mensen allerlei zaken kunnen negeren tot het moment dat ze relevant worden? Hoe doen we dat en hoe implementeren we dat in een AI-systeem? We weten het domweg niet.'

'De robot als gezelschap is een gênante vertoning'

Max Welling, hoogleraar *machine learning* (UvA) formuleert het in zijn *Over leven met kunstmatige intelligentie* nog wat algemener: de belangrijkste tekortkoming van AI is dat ze nog niet genoeg van de wereld begrijpt om in geheel nieuwe situaties toch de juiste beslissing te nemen. 'Je kunt een mens één enkel voorbeeld van een nieuw object laten zien dat hij nog nooit eerder heeft gezien en hij herkent een volgend exemplaar van dit nieuwe object daarna feilloos.' Mensen plaatsen zo'n nieuw voorwerp in een 'ocean van achtergrondkennis over de wereld zodat we het onmiddellijk een plek kunnen geven'. AI kan dat (nog?) niet.

4

Met welke valkuilen moet je bij het toepassen van AI rekening houden?

Ten eerste: vertekening. De meeste zelflerende AI-systemen zijn namelijk getraind met een door de ontwikkelaar geselecteerde verzameling data. Alleen al de selectie van die data zorgt onvermijdelijk voor een vertekening van de resultaten (de beruchte selectiebias). Als het systeem een uitspraak moet doen op basis van andere data (bijvoorbeeld van patiënten met net iets andere kenmerken), dan zullen de prestaties minder goed of in ieder geval anders zijn, in termen van prognose, diagnostiek of behandeling. Verder is het goed te beseffen dat er voor tal van klinische syndromen – denk bijvoorbeeld aan hartfalen of delier – in de internationale literatuur geen consensus bestaat over de criteria waaraan ze moeten voldoen en dus ook niet over het soort data, labels of feedback waarmee je het AI-systeem voedt.

Tweede punt van zorg: correlatie is geen causatie. Als je met *machine learning* verbanden vindt tussen bepaalde patiëntkarakteristieken en behandeluitkomsten betreft het altijd correlaties. Die zijn daarom niet direct te vertalen naar de kliniek, hooguit zijn ze stof voor hypothesevorming en dus voor nader onderzoek naar mogelijk causale mechanismen. Dat raakt aan punt drie: als artsen twijfels hebben over de toepassing van AI in de klinische praktijk, komt dat, zeggen ze, omdat ze de algoritmes niet doorgronden. Pim

Haselager snapt die onrust wel: 'Wat er tussen input en output gebeurt in een neurale netwerk, ook al hebben we het zelf in elkaar gezet, dat is lang niet altijd duidelijk. Wat daar binnenin gebeurt, is een vaag soort kennisrepresentatie. Het verschilt eigenlijk niet heel erg van hoe onze hersenen functioneren. Je krijgt een stimulus binnen op je retina, niet meer dan een tweedimensionaal plaatje, dat gaat naar de visuele cortex, waar de informatie nog redelijk herkenbaar blijft, maar daarna wordt de gerepresenteerde informatie steeds abstracter, vager en moeilijker traceerbaar, tot het richting motorische cortex langzaam weer "in beeld komt". Maar je mag aan de transparantie van systemen die we zelf maken natuurlijk andere eisen stellen dan aan ons geëvolueerde brein, dus daar moeten we wel aan werken.'

Max Welling vindt die 'blackbox' niet per se een beperking van zelflerende systemen. Als je een AI-systeem bijvoorbeeld wilt leren om met hulp van MRI-beelden de ziekte van Alzheimer op te sporen, kun je op één kenmerk testen, zoals de grootte van de hippocampus, maar ook op andere, subtielere patronen. Dat maakt, volgens hem, de diagnose sterker, maar inderdaad ook ondoorzichtiger en moeilijker te herleiden. Maar dat is bij menselijke beslissingen in wezen niet anders. Een dokter weet vaak ook niet precies waar hij in eerste instantie een diagnose op heeft gebaseerd. Op intuïtie soms, of op vage signalen als het lichtelijk kromlopen van een patiënt.

Toch is het een probleem als je verantwoording moet afleggen over je klinische besluiten, stelt Haselager. 'Daarom is een flink deel van het vakgebied nu doende met *explainable AI* (XAI): hoe kun je een basaal antwoord krijgen op de vraag waarom het algoritme met een bepaalde suggestie komt? Als burger zou ik bijvoorbeeld willen weten of het mijn huidskleur was waarom ik een hypotheek al dan niet gekregen heb. Je kunt dan bijvoorbeeld de input van het systeem variëren, om te zien wat voor gevolgen dat heeft. Maar een expert heeft andere vragen waar weer andere mechanismen een antwoord op moeten geven. XAI vormt dus een complex vakgebied, want het is een AI-probleem dat vaak vraagt om een AI-oplossing.'

Dat maakt toezicht – de vierde kwestie – niet gemakkelijk. Maar het is wel noodzakelijk,



vindt de Britse wiskundige en AI-deskundige Hannah Fry, auteur van *Algoritmes aan de macht* (zie Media & Cultuur op blz. 25). Er moet volgens haar een toezichthouder komen die niet alleen het intellectuele eigendom van algoritmes beschermt, maar die er ook op toeziet dat de (maatschappelijke) voordelen van het gebruik ervan opwegen tegen de nadelen – naar het voorbeeld van de toelating van nieuwe geneesmiddelen. ‘Je moet de chemische samenstelling van de verkochte middelen bestuderen en controleren of ze werkelijk doen wat ze beloven. Precies hetzelfde geldt voor algoritmes. Je kunt van de gemiddelde burger niet verwachten dat die zich in Bayesiaanse kansberekening of datamodelle-ring verdiept, of genoeg van computers weet om de achterliggende code op te zoeken en te analyseren of die netjes te werk gaat. Dat is

‘De overheid moet algoritmes voor ons doorlichten en controleren’

niet reëel. Net zoals je niet kunt verwachten dat datawetenschappers zelf een gedragscode opstellen en zich daar dan ook allemaal aan houden. Dat moet echt van de overheid komen. De overheid moet die algoritmes voor ons doorlichten en controleren of ze doen wat ze beloven, en wel op zo’n manier dat het niet schadelijk is voor mensen.’

Geen sinecure, aldus een groep verontruste onderzoekers in Science begin van dit jaar, want anders dan geneesmiddelen of medische hulpmiddelen zijn algoritmes naar hun aard geen statische producten. Naarmate ze aan meer data worden blootgesteld verandert immers de kwaliteit van hun prestatie. Daarom zou een organisatie als de FDA ‘rigoureuze audits’ moeten uitvoeren na toelating – een soort ‘postmarketing surveillance’, zoals dat ook bij medicijnen gebruikelijk is. ■

web

De bronnen en meer over dit onderwerp vindt u onder dit artikel op medischcontact.nl/artikelen.

VELDWERK

DE FORENSISCH ARTS



JEROEN TIMMERMAN werkt bij de GGD in Amsterdam

DNA-dilemma

Vandaag moet ik DNA afnemen bij iemand die verdacht wordt van een gewapende overval op een oude winkelier. Het slachtoffer is met een pistool tegen het hoofd geslagen, gevallen en heeft een arm gebroken. Een ernstig delict. De verdachte weigert DNA af te geven en nu is dat ‘gevorderd’ door de officier van justitie. Volgens de Nederlandse wet moet een arts dat dan doen.

De politie heeft het goed voorbereid, een bijstandsteam staat klaar met schilden en helmen om de man zo nodig tegen de grond te werken. Ik loop de cel in en vraag of hij wil meewerken. Hij zegt nee.

Dilemma, ik heb dit nog niet eerder meegemaakt. Wat te doen?

Ik kan als arts alleen iets tegen de wil van de patiënt doen als die wilsonbekwaam is en de behandeling in diens belang is. Een delier, een psychose, verminderd bewustzijn, enzovoort.

Maar iemand laten weggemen met een weigering bij zo’n ernstig delict, is ook niet juist. Het zal je vader of

EEN
BIJSTANDS-
TEAM STAAT
KLAAR MET
SCHILDEN EN
HELMEN

opa maar zijn. En een DNA-profiel is welbeschouwd een omhooggevallen vingerafdruk. Als hij onschuldig is, kan hij toch gewoon meewerken? Dit soort dilemma’s horen bij forensische geneeskunde; bloedproeven na verkeersincidenten zijn ook een voorbeeld. Je kunt dit werk ook alleen

doen als je het eens bent met de rechtsstaat waarin je werkt. Ik zou geen forensisch arts willen zijn in China of Pakistan.

De man ziet de schilden op zich afkomen en zegt ineens toch te willen meewerken. Een gevalletje straatcommunicatie: wie knippert het eerst met zijn ogen? Ik accepteer dit nu als toestemming en neem wangslimvlies af. Hij blijkt later schuldig en zit een lange straf uit. Mijn ambivalentie blijft.