

VAN DATA TOT PATIËNT

Ontwikkeling kunstmatige intelligentie vergt lange adem

Kunstmatige intelligentie rukt op in de samenleving, maar in de zorg gaat het langzaam. Een model dat zijn oorsprong kent in de ruimtevaartindustrie, laat ontwikkelaars van zorginnovaties de juiste stappen zetten.

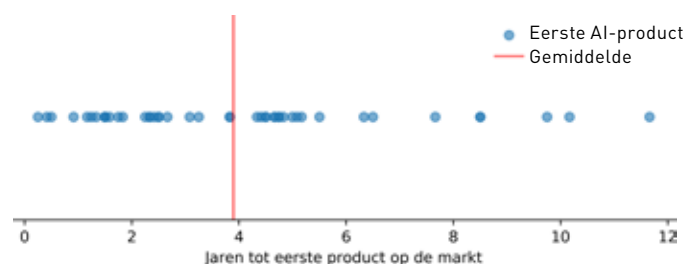
Er is geen medisch congres, vakblad of strategisch plan in de gezondheidszorg waar het onderwerp niet aan bod komt: kunstmatige intelligentie (artificial intelligence: AI, zie kader blz. 9). Het jaarlijkse aantal publicaties op PubMed over AI is tussen 2000 en 2020 ruim vertienvoudigd. Naast onderzoek wordt ook de eerste vertaalslag naar producten gemaakt, met name in het vooroplopende gebied van de medische beeldvorming, waar intussen al meer dan 150 AI-softwarepakketten voor de Europese markt beschikbaar zijn. Denk aan algoritmen die automatisch longnodules detecteren of hersengebieden kwantificeren ten behoeve van de diagnose van alzheimer. Schattingen over de wereldwijde markt voor AI in de gezondheidszorg in 2026 lopen uiteen van 6 tot 44 miljard USD. Met zulke bedragen is ook de verwachting dat de impact die AI op de gezondheidszorg zal hebben, enorm wordt. Naast een voorspelde kostenbesparing door versnelde diagnostiek, personalized medicine en automatisering, verwacht het merendeel van de artsen dat AI een positieve invloed op het zorgsysteem gaat hebben.^{1,2} Ook in Nederland is er veel aandacht voor de gevolgen van AI op ons zorgsysteem. Het ministerie van Volksgezondheid, Welzijn en Sport heeft het programma 'Waardevolle AI voor gezondheid' gelanceerd, waarin de gezondheidszorg geholpen wordt om de potentie van AI om te zetten naar waarde voor patiënt, zorgverlener en burger. Volgens dit programma waren er vorig jaar vierhonderd AI-toepassingen in Nederland in ontwikkeling, variërend van pilotfase tot eindproduct op de markt. Van deze producten heeft 39 procent een CE-markering, waarmee het product voldoet aan wettelijke EU-eisen op het gebied van veiligheid, gezondheid en milieu. Een studie binnen de radiologie laat zien

dat CE-certificering nog niet veel zegt over de effectiviteit en impact van het product op de zorg. Van de honderd commercieel beschikbare producten uit dit onderzoek, bleken er slechts 36 gebaseerd op wetenschappelijke studies over het functioneren en de doelmatigheid van de software.³ De uitgevoerde studies waren voornamelijk retrospectief en gericht op stand-alone performance van AI-software. Bewijsvoering voor de daadwerkelijke toegevoegde waarde van deze AI op het zorgproces is nog minimaal.

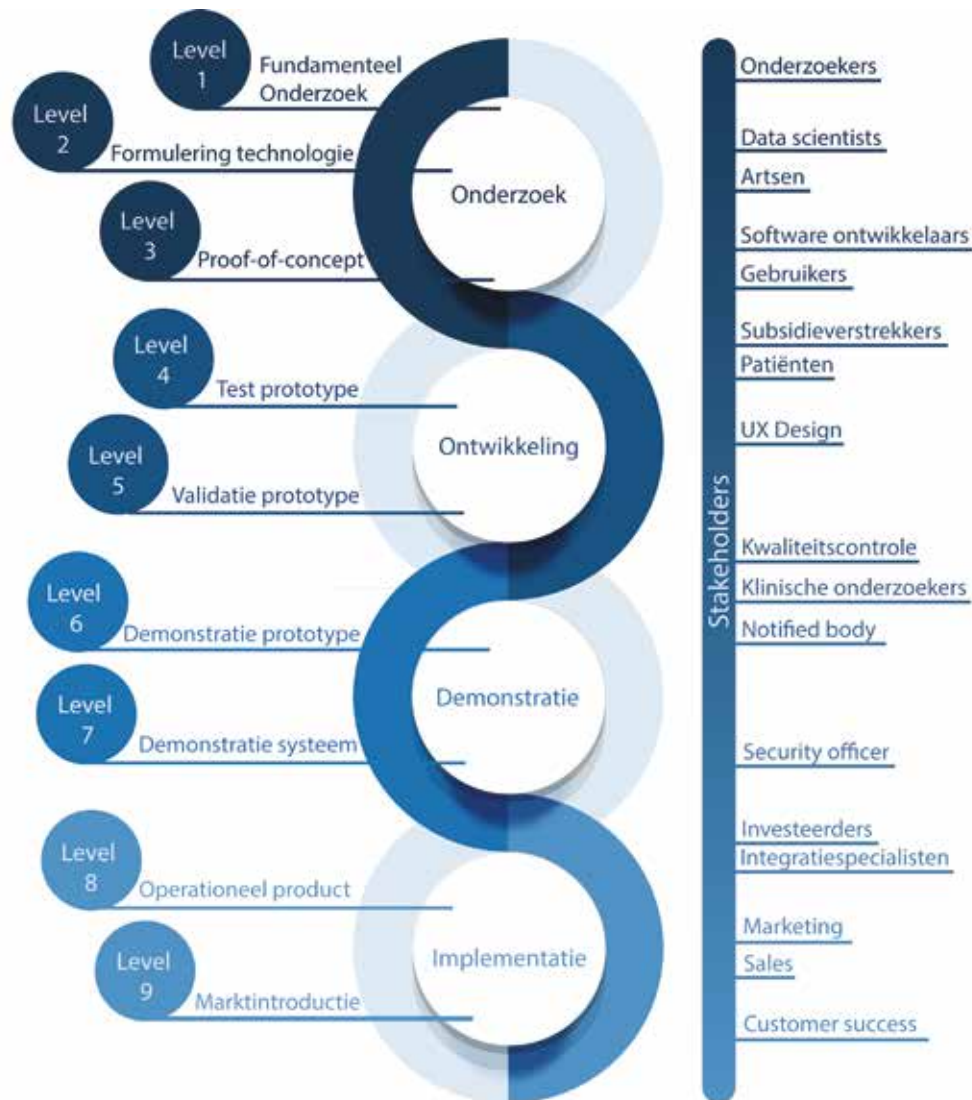
Klinische praktijk

Ondanks alle aandacht voor kunstmatige intelligentie en investeringen in zorginnovaties blijven de klinische implementatie en zichtbaarheid in de gezondheidszorg uit. Hoe kan dat? Om maar met de deur in huis te vallen: de zorgsector staat niet bekend om zijn snelle adoptie van innovatie. Dit heeft tal van redenen, waaronder de benodigde validatie en regulatie die gelden voor medische technologie. En gelukkig maar, want de gevolgen van software die verkeerde aanbevelingen doet in de gezondheidszorg zijn van een ander kaliber dan het Netflix-algoritme. Voor een innovatie in de zorg duurt het gemiddeld zeventien jaar om de route van ontwikkeling tot klinische praktijk af te leggen.⁴ De grootste ontwikkelingen omtrent AI hebben zich afgespeeld in de afgelopen vijf à tien jaar, dus dat we dit nog niet terugzien in de huidige klinische praktijk is eigenlijk niet zo vreemd.

1. TIJD TUSSEN BEDRIJFSOPRICHTING EN EERSTE RADIOLOGIE-AI-PRODUCT OP DE MARKT



2: TECHNICAL READINESS LEVELS, EEN GENERIEK MODEL DAT VERSCHILLENDE STAPPEN BESCHRIJFT OM VAN IDEE TOT IMPLEMENTATIE VAN TECHNOLOGIE TE KOMEN



Als we verder inzoomen op wat er nodig is om van een idee tot een softwarepakket te komen dat gebruikmaakt van AI en in de klinische praktijk functioneert, zien we dat er enkele vereisten zijn. Het Technical Readiness Level (TRL)-model, dat zijn oorsprong kent in de ruimtevaartindustrie in de jaren zeventig, illustreert dit goed.^{5,6} Het TRL-model beschrijft welke negen sequentiële levels doorlopen dienen te worden voordat een product op de markt gebracht kan worden: van basisonderzoek (level 1) tot de commerciële inzet (level 9).

Voorspellend model

Voordat een AI-product bij level 9 komt, gaat er heel wat aan vooraf. De meeste innovaties beginnen bij onderzoek in TRL 1, 2

en worden als proof-of-concept ontwikkeld in TRL 3. Dit kan in zowel het ziekenhuis of academische setting zijn, maar ook in bijvoorbeeld een research & development-afdeling van een bedrijf. Dit wordt voornamelijk gedaan door onderzoekers en data scientists. Daarnaast moet ook de gebruiker hier sturing geven zodat zinvolle toepassingen worden ontwikkeld. Vervolgens wordt het voorspellende model in een laboratoriumomgeving gevalideerd met retrospectieve data (TRL 4 en TRL 5), bij voorkeur data die onafhankelijk zijn van de data gebruikt in TRL 1, 2 en 3. Dit geeft een eerste indicatie van het functioneren van het algoritme. Maar aan een algoritme op zichzelf heeft een arts weinig; aan een softwarepakket wel. Het product wordt verder ontwikkeld in TRL 6 en 7, waar het

product getest wordt in een klinisch relevante omgeving. Hierbij gaat het ook om ontwikkeling van de user-interface, koppeling van de databronnen, anonimisering van de data en de veilige opslag van gegevens lokaal of in de cloud. Een vereiste is ook dat het systeem robuust is, zodat gelijksoortige data gelijksoortige antwoorden leveren en er alleen voorspellingen gedaan worden als de data ook passen bij de data waar het model voor is getraind. Tegelijkertijd moeten de ontwikkeling en validatie gedocumenteerd worden, zodat het product voldoet aan wettelijke EU-eisen (CE-markering). De hoge kosten van de ontwikkeling van het prototype kunnen een reden zijn om de onderzoeksomgeving te verlaten en een bedrijf op te richten. Wellicht stappen in deze fase ook externe financiers in.

In TRL 8 en 9 worden alle eindjes aan elkaar geknoopt tot een systeem dat getoetst kan worden in de uiteindelijke omgeving. Dit is het moment voor gebruik van de AI-software in een gecontroleerde omgeving middels prospectieve studies of simulatiestudies. Na CE-certificering mag het product officieel verkocht worden in Europa.

De zorgsector staat niet bekend om zijn snelle adoptie van innovatie

Wanneer het product commercieel beschikbaar is, kunnen ook de implementatie-engineers, marketeers en salesvertegenwoordigers aan de slag. In deze fase worden, idealiter met behulp van continue monitoring en evaluatie van de AI-software, de kinderziekten verholpen en kan ook de implementatie worden opgeschaald.

Als we naar de AI-producten binnen de radiologie kijken, duurt het gemiddeld vier jaar om van oprichting van het bedrijf (groot TRL 5) tot de marktintroductie (TRL 9) van het eerste product te komen. Hiervoor is vaak al een aantal jaar onderzoek gedaan. Kortom, een AI-algoritme van computer-to-bedside brengen vergt een lange adem.

Vroege evaluaties

Toch zijn er mogelijkheden om AI sneller bij de patiënt te krijgen. Zo kan men met de methodiek Early Health Technology Assessment vroegtijdig inzicht krijgen in de te verwachten consequenties voor zowel de kosten als gezondheidsuitkomsten. Door al vroeg te evalueren of de technologie de potentie heeft kosteneffectief te zijn, kunnen de aandacht en middelen gaan naar die

WAT IS KUNSTMATIGE INTELLIGENTIE?

Kunstmatige intelligentie (AI) omvat het domein van algoritmen en modellen die tekenen vertonen van intelligentie en patroonherkenning door training op data uit het verleden. Voorbeelden van AI-toepassingen in de zorg zijn het gebruik van algoritmen voor detectie van maligniteiten op mammografie, automatische diagnosticering van atriumfibrilleren op basis van het ecg, of voor een patiëntspecifieke inschatting van de kans dat immunotherapie gaat aanslaan. Met name de ontwikkeling van deep learning, een specifiek onderdeel van AI, heeft veel nieuwe mogelijkheden gebracht binnen de medische beeldvorming.

innovaties die daadwerkelijk waarde toevoegen of kan de innovatie gestuurd worden. Het adagium luidt *fail fast, fail cheap*, zodat je niet pas bij TRL 9 ontdekt dat de kosten van de software eigenlijk niet uit kunnen en vergoeding van de zorgverzekeraar uitgesloten is. Health Innovation Netherlands is een initiatief dat dit soort vroege evaluaties faciliteert met behulp van rondetafelgesprekken.⁷

Verder moet de focus liggen op het verkrijgen van (transparante) validatieresultaten. Het is van belang dat aangetoond wordt dat de AI-modellen goed functioneren en ook waarde toevoegen in andere klinische omstandigheden. Dit kan zowel geografisch, demografisch als economisch zijn. De Tripod-checklist voor modellering en validatie van predictiemodellen wordt gebruikt als leidraad voor modelontwikkeling.⁸

Tot slot moeten ziekenhuizen zich realiseren dat er meer nodig is dan een geschikte use-case en een datascientist om tot een veilig AI-product te komen dat waarde toevoegt aan de dagelijkse klinische praktijk. Het TRL-model toont dat overduidelijk aan. Kortom, diversiteit in kennis en kunde is hard nodig. Daarnaast zijn wederzijds vertrouwen en realistische verwachtingen over functionaliteit, benodigde capaciteit en het tijdspad essentieel om tot breed gedragen implementatie van AI te komen. Het allerbelangrijkste is geduld; geduld om producten degelijk te ontwikkelen en valideren, te leren samenwerken met de technologie, en te zorgen voor veilige en waardevolle implementaties van AI. Het belangrijkste in de medische wereld blijft *first do no harm*, dit geldt ook voor de toepassing van AI. ■

contact

kicky.vanleeuwen@radboudumc.nl
cc: redactie@medischcontact.nl

web

De voetnoten vindt u onder dit artikel op medischcontact.nl.