

De computer bepaalt vraagt om waakzaamheid

'Computer says no'

Deze uitdrukking – vooral bekend geworden door de hilarische sketches uit de comedyserie *Little Britain* – slaat op situaties waarin mensen of organisaties hun handelen volledig baseren op de informatie die het computersysteem hen verschaft, en waarbij het gezonde verstand en elke vorm van redelijkheid zijn verdwenen. De absurde situaties die aan de creatieve geesten van de scriptschrijvers van *Little Britain* zijn ontsproten, lijken steeds vaker realiteit te worden.

Het fenomeen dat mensen niet verder kunnen, mogen of willen kijken dan hun beeldscherm diep is, kun je overal tegenkomen. Ik heb zelf al eens mogen meemaken dat mij de toegang tot een vliegtuig werd geweigerd, omdat ik volgens de computer al aan boord zat. En ik - wapperend met tickets, paspoort en andere papieren - de man achter de balie maar niet tot het inzicht kon brengen dat de informatie op zijn beeldscherm onjuist móest zijn, aangezien ik toch echt in levende lijve voor hem stond. Pas toen het gemor in de rij achter mij aanzwol werd door de stoïcijnse functionaris ondersteuning ingeroepen. Na nogmaals veel heen en weer gepraat - en vluchtvertraging - mocht ik uiteindelijk mee.

Het blijft niet altijd bij een klucht. Zo klaagden de drie grootste hypotheekbemiddelaars afgelopen juni bij BNR Nieuwsradio dat onnodig veel mensen geen hypotheek krijgen omdat de computersystemen van geldverstrekkers geen rekening houden met persoonlijke situaties. Dit onderwerp kreeg nadien een verdieping door media-aandacht voor het Bureau Krediet Registratie (BKR). Geldschietters moeten het BKR verplicht raadplegen. Begrijpelijk, want het is van maatschappelijk belang dat wordt getoetst of consumenten een krediet kunnen dragen (denk maar aan de kredietcrisis, die zijn wortels mede had in de misstanden op de Amerikaanse hypotheekmarkt). Het BKR ligt echter onder vuur omdat het nagenoeg onmogelijk blijkt om een negatieve registratie te wijzigen, ook al is deze futiel of achterhaald. Dat kan er toe leiden dat een rekening van een paar tientjes, die twee dagen te laat is betaald als gevolg van verkeerd bezorgde post, resulteert in een negatieve BKR-registratie, die jaren later verrassenderwijze leidt tot afwijzing van een hypotheekaanvraag. Het BKR wijst kritiek van de hand, aangezien kredietverleners zich zouden (moeten) baseren op het volledi-

ge plaatje van de consument (dus de gezinssituatie qua bezit, inkomen, spaargeld en dergelijke). De Vereniging Eigen Huis stelt dat hypotheekverstrekkers daar in de praktijk geen zin in of tijd voor hebben. Hoe dan ook is de feitelijke situatie dat een bredere beoordeling in de regel achterwege blijft. De interne procedures van banken dicteren dat er geen hypotheek wordt verstrekt aan personen met de bewuste BKR-codering. De keuze om op deze wijze met informatie om te gaan draait voor veel mensen daadwerkelijk uit op: Computer says no.

Computergestuurde besluitvorming dringt door in steeds meer domeinen van de samenleving. Een overweging die hierbij meespeelt is dat computergestuurde besluitvorming het voordeel zou hebben van objectiviteit, in contrast met de subjectiviteit die inherent is aan menselijke besluitvorming. Zo laat een recruiter van bureau YoungCapital in NRC¹ weten erg enthousiast te zijn over de succesindicatie die computeralgoritmes genereren over individuele sollicitanten. Hij stelt dat hij nu in één oogopslag kan zien welke sollicitanten het meest geschikt zijn voor een bepaalde functie en bijna blind op de voorspellingen te vertrouwen. Dit is niet onomstreden. Zo uit Van Luijk, selectiepsycholoog bij bureau LTP, principiële bezwaren². Selectie via algoritmes resulteert volgens hem in conservatisme. Immers, de data hebben betrekking op het verleden; het algoritme kijkt achterom, naar vroegere successen. Organisaties krijgen hierdoor meer van wat ze al hebben. Terwijl er juist geen behoefte is aan klonen, maar aan daadwerkelijk nieuwe mensen met veranderkracht, aldus Van Luijk. Hoogleraar kunstmatige intelligentie Van Harmelen stipt een verwant punt aan, namelijk de kwaliteit van de data-invoer. Het algoritme doet voorspellingen op basis van eerder beslisgedrag van recruiters. Maar hoe en op basis waarvan deze eerdere beslissingen tot stand zijn gekomen is niet te achterhalen. Daar liggen allerlei - al



WFZ
Waarborgfonds
voet de Zorgsector

Herman Bellers,
directeur WFZ



dan niet juiste - aannames, veronderstellingen en (voor) oordelen aan ten grondslag. Vanzelfsprekend is het zo dat wat er aan input het systeem in gaat, er vervolgens als output net zo hard uit komt. Terwijl de gebruiker geen zicht heeft op de onderliggende data en de al dan niet aanwezige mankementen daarin.

Duidelijk moge zijn dat het vermeende voordeel van objectiviteit als sneeuw voor de zon verdwijnt in alle gevallen waarin de menselijke subjectiviteit in de data-invoer zit verweven. Een illustratie hiervan is de schoonheidswedstrijd op basis van kunstmatige intelligentie waarover het FD vorig jaar verslag deed³. Juist om subjectieve discussies over 'mooi' en 'lelijk' te vermijden werd een computer ingeschakeld om objectief de symmetrie van het gezicht, de gelijkmatigheid van de huid en andere relevant geachte variabelen in samenhang te rangorden. Bij de uitslag bleek de top 40 geheel blank te zijn, terwijl het merendeel van de ruim zesduizend deelnemers een Afrikaans of Aziatisch uiterlijk had. Nader onderzoek wees uit dat het computeralgoritme was gebouwd en 'getraind' op basis van foto's van vooral blanke modellen. Een ernstiger voorbeeld van bias-risico is dat in de VS computers worden ingezet om te voorspellen of criminelen later opnieuw de fout in zullen gaan. De rechter geeft aan de hand van dit risicoprofiel een hogere of lagere straf. Niet-blanken blijken systematisch een hoger risico en daardoor hogere straffen toegekend te krijgen. Analyse van daadwerkelijke recidivecijfers geeft echter grond voor serieuze twijfel aan de rassenneutraliteit van de data en beslisregels⁴.

ICT-ontwikkelingen kunnen voor de mensheid een zegen zijn. Maar door een misplaats geloof in de on-

feilbaarheid van moderne computertechnologie, al dan niet in combinatie met gemakzucht of kosten-baten overwegingen, kan die zegen soms in een vloek veranderen. Vooral wanneer de onderliggende data en beslisregels een black box zijn is het oppassen geba-zen. Ook met betrekking tot de digitalisering van de zorgverlening lijkt mij dit een aandachtspunt. Zo berichtte Zorgvisie medio juli over een slim zelflerend computersysteem dat meeleeft in het EPD en automatisch signalen geeft aan artsen en verpleegkundigen. Mij viel de opmerking op dat met het systeem '...patronen kunnen worden gedestilleerd die voor een persoon niet te achterhalen zijn, zelfs als hij voldoende tijd zou hebben om een dossier helemaal door te spitten'. De bezwering dat het systeem 'zelf geen beslissingen neemt' en 'de specialist of verpleegkundige altijd de supervisie houdt over het systeem' lijkt mij in dit licht tamelijk gratuit.

De gedachte dat een medische ingreep aan mijn lijf en leden - of het nalaten ervan - zou worden bepaald op basis van computeroutput, waarvan de totstandkoming zelfs voor de behandelaar niet te traceren is, ervaar ik vooralsnog niet als positief voor de gemoedsrust. Ongefundeerde digitaliserings-sceptis? Mogelijk. Maar de ontwikkelingen en bevindingen in andere domeinen - zie de voorbeelden - geven naar mijn mening enige reden voor waakzaamheid. ◆

¹ NRC, 27 juli 2017

² NRC, 2 augustus 2017

³ FD, 10 november 2016

⁴ Terzijde: in het tijdschrift 'Justitiële verkenningen' werd al in 1992 (jrg. 18, nr. 8) ingegaan op de discriminatie naar ras die inherent is aan het hanteren van bepaalde recidive-gecorrleerde factoren, alsmede het nagenoeg ontbreken van voorspellende waarde van recidive-gerelateerde variabelen.