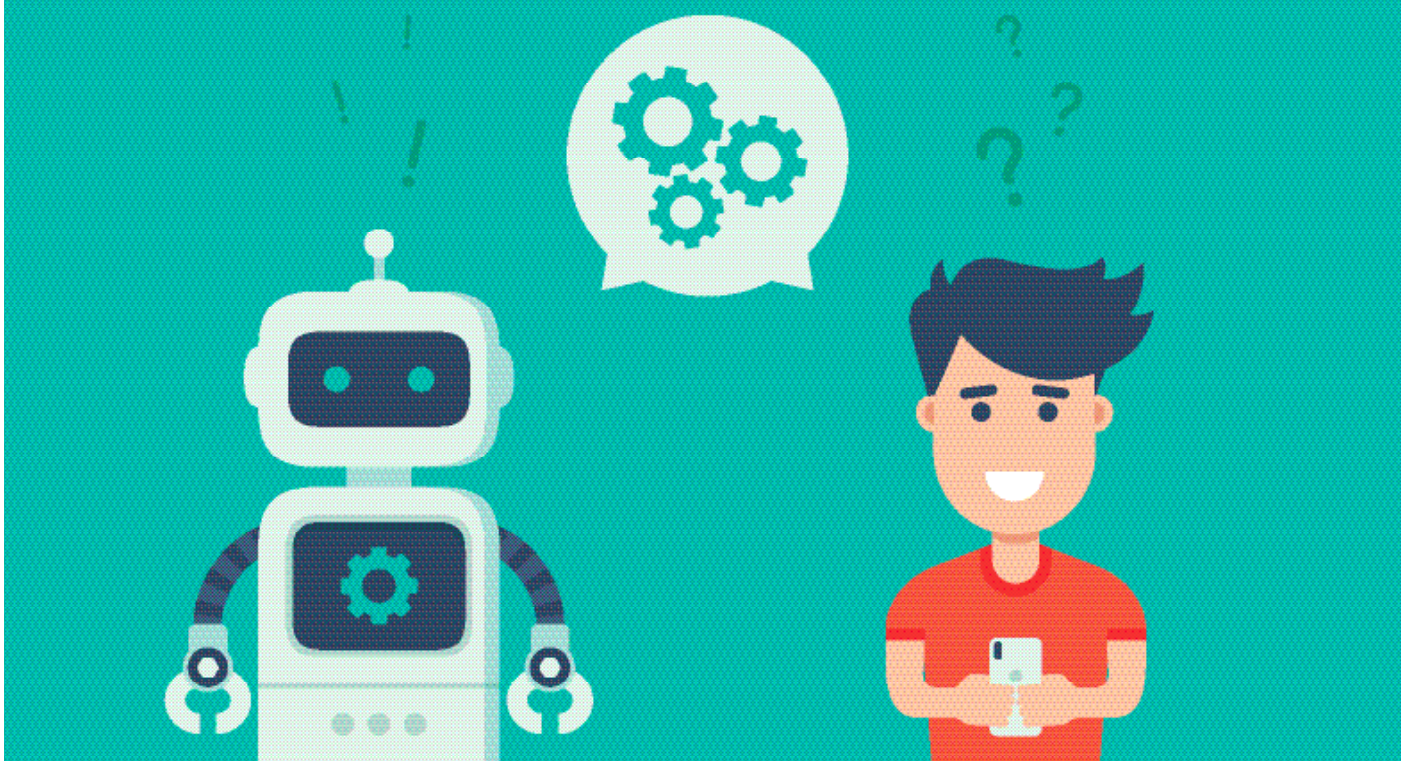


AI als gesprekspartner kan meer ruimte voor menselijke maat bieden

ARTIFICIAL INTELLIGENCE: BIAS UITBANNEN OF IN GESPREK MET AI?



BIJ 'ARTIFICIAL INTELLIGENCE EN ETHIEK' DENK JE AL GAUW AAN HET UITBANNEN VAN BIAS. MAAR WAT NU ALS WE AI OPVATTEN ALS EEN GESPREKSPARTNER DIE, NET ALS MENSELIJKE GESPREKSPARTNERS, VOORINGENOMEN IS? WAT KAN DIT PERSPECTIEF BIJDRAGEN AAN HET BEHOUD VAN DE MENSELIJKE MAAT IN EEN WERELD WAARIN AI STEEDS MEER AANWEZIG IS?

door Henriëtte Joosten beeld Shutterstock

DE TERM 'DATAÏSME' WERD IN 2013 VOOR HET EERST GEBRUIKT DOOR JOURNALIST DAVID BROOKS. Met deze term beschreef hij de 'filosofie' of het rotsvaste vertrouwen dat big data onze wereld zullen verbeteren. Hoe meer data, hoe beter. De computer, gevoed door bergen data, weet vele malen beter dan mensen wat werkt en wat goed voor ons is. In 2021 is het dataïsme springlevend. Of het nu gaat om screening van kanker, zelfrijdende auto's of het voorspellen van solvabiliteit: AI leidt vanzelfsprekend tot objectievere, en dus betere beslissingen dan wanneer mensen aan zet zijn. En de toekomst kan alleen maar beter worden; er worden immers steeds meer data vastgelegd. In haar boek 'Frictie. Ethiek in tijden van dataïsme' (2020) constateert Miriam Rasch: "De voorspellende kracht van data, hun werking en hun waarde, lijkt tegenwoordig boven alle twijfel verheven te zijn."

ETHIEK IN TIJDEN VAN DATAÏSME

Natuurlijk zijn er ethische kwesties. Bias – of op z'n Nederlands: vooringenomenheid, vooroordeel, vertekening – is een belangrijke. Ontwerpers van algoritmes gebruiken (vaak ongemerkt) hun normen en waarden om AI van kleuring te voorzien. Dat geldt ook voor de datasets waarmee AI wordt getraind. Zo herkent gezichtsherkenningsoftware gezichten van mannen veel beter dan van vrouwen. En hoe donkerder de huid, hoe minder herkenning. Volgens Joy Buolamwini, onderzoeker bij het MIT Media Lab, komt dit doordat de programmeurs vooral mannen met een lichte huid zijn. Het trainingsmateriaal weerspiegelt hun wereld. Er wordt niet alleen gewaarschuwd voor de gevolgen van bias. Allerlei tools en checklists zien het licht om bias aan te pakken. De overheid doet haar best om toezicht op de onwenselijke impact van algoritmes te institutionaliseren. Zo hopen we te kunnen vasthouden aan het

“Eén keer
googelen en de
bankmedewerker
had geweten wie
hij was”

aantrekkelijke idee dat AI – als er maar genoeg data zijn – ons vertelt wat goed voor ons is. Waar het dataïsme hoogtij viert, richt ethiek zich vooral op het uitbannen van bias.

RUIMTE VOOR DE MENSELIJKE MAAT

Wat AI als 'goed' bestempelt, is niet voor iedereen goed. In onze samenleving zijn er vele vormen van 'goed'. En hoeveel ruimte is er voor de veelbesproken 'menselijke maat' wanneer het dataïsme domineert? Deze maat vergt juist lef om je niet al te zeer vast te laten leggen door cijfers en protocollen, maar deze telkens weer naar eigen inzicht te interpreteren. Daar komt bij dat 'het goede' verandert. Ondernemers willen steeds vaker sociaal-maatschappelijke en ecologische waarden creëren. Deze beweging richting meervoudige waardecreatie vergt (mentale) ruimte voor overleg, nieuwe relaties en gezamenlijke betekenisgeving. Zijn er voldoende 'vrijplaatsen' waarbinnen het goede leven en de menselijke maat gestalte kunnen krijgen in een wereld vol AI?

AI ALS GESPREKSPARTNER

Je kunt AI ook als gesprekspartner benaderen. Net als mensen wordt deze gesprekspartner gekenmerkt door 'voor-

ingenomenheid'. We zijn altijd in de wereld vanuit een zekere vooringenomenheid, of gerichtheid. Een gesprek kan ook pas ontstaan vanuit deze kleuring. Het is een gesprekspartner die je bij de gang van zaken betreft en wiens vooringenomenheid ook onderwerp van gesprek is. Net als de bias van mensen. Wat brengt dit perspectief in beeld? Laten we eens kijken naar een voorbeeld uit de medische wereld.

AI EN DE DIAGNOSE VAN TUMOREN

Algoritmes zijn uitstekend in staat om op basis van MRI-scans de aanwezigheid en zelfs de agressiviteit van prostaat-tumoren te berekenen. Normaliter is de beoordeling van de MRI-beelden mensenwerk. Maar algoritmes blijken het prima te kunnen overnemen. Een groot voordeel is dat de verschillen in beoordeling tussen radiologen (lees: menselijke bias) worden voorkomen omdat deze techniek de beoordeling standaardiseert en automatiseert. En dan te bedenken dat deze toepassing nog maar in de kinderschoenen staat.

WAT WE VERLIEZEN BIJ EEN DATAÏSTISCH VERTROUWEN IN AI

Het werk van radiologen lijkt op de tocht te staan. Maar er gebeurt meer: het huidige diagnostische proces als herhaalde betekenisgeving staat op het spel. Tijdens dit proces geven behandelaar, patiënt en naasten en andere specialisten betekenis aan een veelheid aan gegevens, waaronder de lichamelijke signalen en de verhalen van de patiënt, hypothesen van specialisten, gegevens in het patiëntendossier en de cijfers en beelden die scans en andere instrumenten opleveren. Vanuit een dataïstisch vertrouwen is het een kleine stap om AI te beschouwen als een objectieve leverancier van onmiddellijk uitvoerbare behandelingsuggesties. Die beweging is al gaande. Artsen vinden het uitdagend om op algoritmes gebaseerde oordelen te weerleggen en

AUTEUR



HENRIËTTA JOOSTEN
is filosoof, bedrijfskundige
en associate lector
Financial Ethics aan de
Hogeschool van Arnhem
en Nijmegen (henrietta.
joosten@han.nl). Zij is lid
van de KNVI.

in de medische praktijk te omarmen, is er meer ruimte voor:

- intuïtie en ervaring van artsen bij het iteratief interpreteren van gegevens;
- de menselijke maat; dat wil zeggen de maat van de patiënt: wat telt voor hem als een goed behandeltraject?
- veranderende visies in de samenleving op ziekte en gezondheid;
- leren van AI: AI bleek bij de diagnose van huidkanker veel beter te scoren dan artsen. Na onderzoek bleek dat artsen hun blik moesten verbreden naar het gebied rondom een tumor;
- de vormgeving van AI: kleur en grootte van de cijfers, de volgorde van resultaten en de kopjes ('second opinion' of 'geconstateerd') beïnvloeden stuk voor stuk hoe mensen de inbreng van AI oppakken. Die beïnvloeding moet ook onderwerp van gesprek zijn.

Nog een voorbeeld.

DE MENSELIJKE MAAT BIJ FRAUDEDETECTIE

Banken zijn bij wet verplicht om actief te monitoren op ongebruikelijke transacties en deze te melden. AI is hierbij onmisbaar. Een collega vertelde mij hierover een mooie anekdote. Een gewaardeerde CEO van een Nederlands miljardenbedrijf is net met pensioen. Hij is een zuinig mens gebleven. Naar de stad gaan hij en zijn vrouw op de fiets en zelfgesmeerde boterhammen gaan mee. Ze wonen in een rijtjeshuis. Uitge-rekend deze man wordt op een dag door een bankmedewerker gebeld. Waarom hij zo'n grote som geld op zijn rekening had staan. Mijn collega zei: "Eén keer googelen en de bankmedewerker had geweten wie hij was." De bank kan haar algoritmes uitbreiden ter voorkoming van dit soort 'ongelukjes'. Maar daarmee blijft een onderliggend probleem, gerelateerd aan het dataïsme, bestaan: medewerkers ervaren weinig ruimte om 'voorbij' AI-aanbevelingen te kijken. Ook is het maar de vraag of er vrije ruimte is voor gezamenlijk overleg over schurende zaken, veranderende inzichten en maatschappelijke verantwoordin-

ging; of de vraag 'wat doet er in ons werk werkelijk toe?' gesteld kan worden. De voormalig CEO werd serieus genomen. Zijn naam verdween al snel van het 'verdachte lijstje'. Bij duizenden burgers van de toeslagaffaire was dat niet het geval. Wat meer vrije ruimte was op zijn plaats geweest.

OPENING

Laten we vooral profiteren van de ongekende mogelijkheden van AI én bias voorkomen. Maar laten we ook erkennen dat bias nooit helemaal uit te bannen is en dat onze houding ten opzichte van AI grote invloed heeft op de ruimte voor de mens als betekenisgever. 

Zie www.agconnect.nl voor de referenties.

**WAT AI ALS
'GOED' BESTEMPELT,
IS NIET VOOR
IEDEREEN GOED**

onafhankelijke diagnoses en behandel-suggesties voor te stellen. Hoeveel ruimte blijft er voor patiënt en arts om samen betekenis te geven aan het door het algoritme gedetecteerde patroon? Mag een patiënt afwijken van de behandel-suggestie omdat hij een andere invulling geeft aan 'kwaliteit van leven' dan het algoritme? Kan/wil een arts nog ingaan tegen het algoritme?

AI ALS GESPREKSPARTNER BIJ DIAGNOSES

AI-toepassingen, net als mensen, zijn geen neutrale, objectieve machines. Door AI als gesprekspartner – met bias –

