

Wolter van der Knaap – De sirenenzang van artificiële intelligentie

Pleidooi voor kritisch taalgebruik rond data

22

Wat denkt u? Hoeveel besluiten in uw leven worden op basis van artificiële intelligentie (AI) voor u genomen? Waarschijnlijk meer dan u denkt. AI heeft wel iets van het karakter van de *Matrix* uit de gelijknamige film uit 1999: overal aanwezig en tegelijk onzichtbaar. Het maakt ons leven aangenaamer, gemakkelijker, zonder dat u zich altijd bewust bent van wat er gebeurt. Maar hoe zit het met het voorrecht om zelf in ons leven besluiten te nemen en daar verantwoordelijk voor te zijn?

Stel: u rijdt 's avonds na een vergadering naar huis. Uw zelflerende thermostaat weet uit uw gekoppelde agenda tot hoe lang de vergadering duurt en heeft inmiddels ook geleerd dat het meestal een half uurtje uitloopt. Onderweg doet zich echter onverwacht een vertraging voor van vijftien minuten wegens reparatiewerkzaamheden aan de snelweg. Geen probleem, uw telefoon geeft het door aan uw thermostaat. Als u thuiskomt, is de temperatuur in de huiskamer aangenaam. U merkt het niet eens, zo natuurlijk voelt het aan! Om te ontspannen besluit u nog even te netflixen. Netflix weet dat u op dit tijdstip eigenlijk nooit langer dan een half uur kijkt, en dus suggereert het u een korte film van 25 minuten. Precies het genre dat u graag ziet na een drukke dag.

Dit is vast geen herkenbaar beeld voor de meeste lezers. Nou ja, niet op alle onderdelen tegelijk. Maar of we het door hebben of niet, anno 2022 maakt AI deel uit van ons leven. Wilt u bewijs? Dat is eenvoudig te leveren. Laten we aannemen dat we allemaal weleens iets op internet opzoeken. Op dit moment gebruik ik zelf twee laptops. De browsers op beide

apparaten zijn identiek (Chrome), maar bevatten natuurlijk wel een eigen zoekgeschiedenis plus bijbehorende *cookies*. Als ik in de ene browser voluit 'Artificial Intelligence' intik in Google, krijg ik bovenaan drie advertenties over de voordelen van AI-toepassingen te zien. De andere browser presenteert er echter maar één. Hoe komt dat? Het antwoord is dat het algoritme van Google aan de hand van gegevens op mijn ene laptop inschat dat alle drie de advertenties passend zijn voor mij en op mijn andere laptop tot de conclusie komt dat één advertentie volstaat.

Ik wil twee belangrijke aspecten van de manier waarop AI zich in onze wereld manifesteert, uitwerken. Het eerste is dat AI ons in het algemeen iets lijkt te beloven: vooruitgang, verbetering, gemak, controle, beheersbaarheid, zeg het maar. Maar wel op één voorwaarde: dat we steeds meer onze besluiten aan AI uitbesteden. Het tweede aspect is subtieler en betreft ons taalgebruik met betrekking tot AI. We lijken de neiging te hebben om over AI te spreken alsof het menselijke eigenschappen heeft. Ik denk dat het belangrijk is om kritisch naar beide aspecten te kijken.

Verwachtingsvolle sfeer

Waar hebben we het precies over? Ik gebruik graag de definitie die Mieke De Ketelaere, programmadirecteur AI bij het Belgische onderzoekscentrum IMEC,¹ geeft in haar recente boek *Mens versus Machine*.² Zij schrijft over het begrip artificieel: 'iets is artificieel als het niet op natuurlijke manier ontstaat of gebeurt, en is een synoniem van kunstmatig.' Dat is duidelijk. De tweede helft van het begrippenpaar is lastiger te pakken. Verschillende disciplines verstaan immers iets anders onder intelligentie. Denk slechts aan de psychologie en de filosofie. De Ketelaere kiest voor een pragmatische definitie: AI betreft volgens haar 'computersystemen die zelfstandig leren, beslissingen nemen en ze ook uitvoeren'. Een goed voorbeeld is het al eerder aangehaalde algoritme van Google. Dat is inderdaad (1) een computersysteem, dat (2) zelfstandig steeds beter leert te onderscheiden wat voor antwoorden een bepaald individu waarschijnlijk wil zien, (3) besluit welke antwoorden wel (of niet!) gegeven

zullen worden, en ten slotte (4) de geselecteerde antwoorden naar het betreffende computerscherm stuurt.

Het eerste dat mij steeds weer opvalt rond AI is de verwachtingsvolle sfeer van vooruitgang die eromheen hangt. Neem bijvoorbeeld de advertenties die mijn browser toont bij het woordenpaar *artificial intelligence*. Het is interessant om naar de allereerste zinnen te kijken van de teksten die op mijn scherm verschijnen. De eerste advertentie stelt mij een vraag: *‘Artificial Intelligence, wat is het en waarom is het zo belangrijk voor mijn organisatie?’* De onderliggende aanname in deze zin is dat AI sowieso ‘belangrijk’ is. De tweede advertentie geeft mij een advies: *‘Haal nog meer waarde uit je data met Machine Learning en Artificial Intelligence.’* Kennelijk ligt er ergens ‘nog meer waarde’ voor het grijpen. En de derde advertentie raadt mij aan kennis te nemen van *‘the possibilities of empathetic AI, and what it means for our future. Can machines do what we can do?’* De strekking is duidelijk. Er ligt een verborgen waarde in jouw data en AI helpt je die waarde te realiseren. Een betere toekomst is binnen handbereik! Waarom zou je dat niet willen?

– Data lijken boven alle twijfel verheven te zijn –

Natuurlijk zijn dit advertentieteksten en dus bedoeld om op een positieve manier aandacht te trekken voor het onderwerp. Maar ze appelleren aan een sentiment dat inmiddels diep in onze culturele bagage lijkt te zitten. De voorspellende kracht van data, hun werking en hun waarde, lijken tegenwoordig boven alle twijfel verheven te zijn, schrijft Miriam Rasch dan ook in haar prijswinnende boek over ethiek in tijden van dataïsme.³

In kleine kringen kan dit geloof in vooruitgang welhaast religieuze vormen aannemen, hoewel zonder (een) God. Zo schrijft de Britse onderzoeker en wetenschapper James Lovelock in zijn boek *Welkom in het Novaceen*⁴ dat mensen in de komende jaren voorbijgestreefd zullen worden door vormen van hyperintelligentie. Na de komst van de mens op aarde beschouwt hij dat als de volgende

stap omhoog in de evolutie. Is dat erg voor de mensheid? Welnee, vindt hij. ‘We moeten zorgen dat de evolutie van al het leven op aarde door blijft gaan opdat we de almaar toenemende gevaren weerstaan die dreigen voor ons en voor Gaia, dat machtige systeem dat al het leven en alle essentiële onderdelen van onze planeet omvat.’ De mensheid is in zijn opvatting slechts een fase in een steeds verdergaande evolutie. Lovelock vindt dat helemaal geen somber beeld, maar is juist trots op wat wij mensen tot stand hebben gebracht.

Dat optimisme wordt niet alleen zichtbaar in advertenties en in speculaties over de toekomst, maar is ook zichtbaar in de serieuze managementliteratuur. De ondertoon is dat er ‘goud’ ligt te wachten voor organisaties en ook dat zij er vooral snel bij moeten zijn. Er is een algoritme-revolutie gaande, betoogt James Taylor, hoogste baas van het Amerikaanse bedrijf Decision Management Solutions.⁵ In toenemende mate zullen het volgens hem algoritmes zijn, en niet mensen, die de beslissingen nemen in een organisatie. Op basis van data zullen de algoritmen ‘inzicht’ ontwikkelen en dat inzicht zullen zij toepassen om zo voordelig, winstgevend, efficiënt en effectief mogelijke besluiten te nemen voor de organisatie.

Slapen op een schat

Dergelijke betogen samen met de overtuiging dat AI toegang biedt tot braakliggend potentieel, leiden tot een gevoel van onrust bij managers van organisaties. Doen we wel genoeg met onze data? Zitten we niet te slapen op een schat terwijl onze concurrent reeds druk bezig is die te ontginnen? En zijn we objectief genoeg met onze analyses? Kunnen we besluiten nog wel overlaten aan per definitie subjectieve medewerkers? Zouden we niet eens een AI-project moeten starten? Al jaren geleden zag ik een cartoon die dat ongeruste gevoel treffend uitbeeldde. Het plaatje liet een reeds vertrekkende trein zien, de coupés gevuld met identiek uitzierende zakenmannen, gekleed in grijze pakken en met aktetassen in de handen. Over het perron kwamen nog twee zulke zakenmannen aanhollen, stapels documenten onder de armen, de een tegen de ander verzuchtend: ‘Ik heb geen idee waar

deze trein heengaat, maar ik ben zó bang hem te missen...’.

Bijna als vanzelf wordt in deze culturele onderstroom de mens, met al zijn feilbaarheid en zijn goede en slechte eigenschappen, gezien als een beperking. Thomas H. Davenport en Jeanne G. Harris, die beiden kunnen bogen op een indrukwekkende staat van dienst in de Amerikaanse wetenschap en het bedrijfsleven, schreven in 2017⁶ dat het knelpunt niet de technologie is, maar de cultuur van organisaties. De waarden en overtuigingen van mensen dus.⁷ Als een organisatie er niet in slaagt om van boven tot beneden ‘*data-driven*’ te worden, zullen de resultaten achterblijven, betogen zij. Met andere woorden, medewerkers in organisaties dienen erop te leren vertrouwen dat grote hoeveelheden data (*big data*) plus de logische analyse daarvan, zonder meer zullen leiden tot de beste besluiten.

Er is een groep mensen in de AI-wereld – Davenport en Harris horen daarbij – die nog een stap verder gaat. Data en data-analyse dienen er dan voor om de WAARHEID of de WERKELIJKHEID in kaart te brengen waarmee, zo stellen deze twee auteurs, voorkomen wordt dat besluiten in organisaties worden genomen om bijvoorbeeld de persoonlijke doelen van managers te bevorderen. Het beeld dat zij schetsen is gebaseerd op een tegenstelling: aan de ene kant bevinden zich data en data-analyse – objectief, zuiver, effectief en rationeel – en aan de andere kant de menselijke betekenisgeving en daaraan verbonden besluitvorming – subjectief, onzuiver, en volgens hen vaak gericht op persoonlijke belangen.

– Mensen wel of niet terugsturen als ze zich aan onze grens melden? –

Deze voorstelling van zaken van Davenport en Harris is volgens mij onjuist. Data en betekenisgeving zijn altijd met elkaar verbonden. Alleen is dat niet altijd even zichtbaar. Data zijn op zichzelf immers volkomen betekenisloos. Data krijgen pas

betekenis als iemand die betekenis daaraan gééft. Dat kan expliciet gebeuren in de vorm van een transparant proces van besluitvorming.

Onzichtbaarder

Het kan ook veel onzichtbaarder, bijvoorbeeld door de betekenisgeving al vooraf in analysesoftware in te bouwen. Een voorbeeld kan dat verhelderen. Hierboven haalde ik de stelling van Davenport en Harris aan dat de logische analyse van *big data* zal leiden tot de ‘beste’ besluiten. Maar wat is ‘best’? De impliciete aanname lijkt meestal te zijn dat de beste besluiten die besluiten zijn die de hoogste winstverwachting hebben. Het is niet zo moeilijk om dat criterium in de analysesoftware in te bouwen, maar het zal duidelijk zijn dat daar een hele wereld van subjectieve, cultuur-geladen waarden in meekomt. Hoezo objectief?

Data op zichzelf beantwoorden helemaal geen vragen. Er dient eerst sprake te zijn van een betekenisgevende norm – dat kan bijvoorbeeld een gewenste kamertemperatuur van 20 graden Celsius zijn – en vervolgens kun je aan de hand van verkregen data inderdaad objectief vaststellen of een gemeten temperatuur onder of boven de norm ligt en welke acties we daaraan wensen te verbinden (kachel aan/uit). Nu is dit een type besluit dat we met een gerust hart aan een thermostaat uitbesteden. Maar wat nu als de omstandigheden complex zijn en tegelijkertijd vergaande gevolgen hebben voor de betrokkenen? Denk eens aan situaties waarin gekozen dient te worden voor wel/niet opereren. Of wel/niet doorgaan met behandelen. Of bepaalde belastingbetalers wel/niet op een lijst van potentiële fraudeurs zetten. Of mensen wel/niet terugsturen als zij zich aan onze grenzen melden. Juist de complexiteit van zulke vragen vraagt voortdurend om doordachte besluitvorming in de concrete situatie door verantwoordelijke mensen van vlees en bloed.

Nogmaals, tussen data en besluit zit altijd betekenisgeving. Of dat geautomatiseerd plaatsvindt of niet, is daarvoor niet relevant. Ik pleit ervoor om dat moment van betekenisgeving zo expliciet en transparant mogelijk te maken. Miriam Rasch duidt dat moment met het woord ‘frictie’, om aan te geven dat

het lastig, vertragend kan voelen. Je moet immers even stilstaan en je afvragen welk besluit je gaat nemen op basis van een concrete analyse. Frictie betekent in de praktijk dat je extra tijd neemt om heel bewust een beslissing te nemen op basis van de analyse die je hebt gemaakt en dat je dus niet zo snel mogelijk doorgaat met actie.

Dat ‘tussenmoment’ is ook het moment dat ethiek een belangrijke rol kan spelen. Soms is die rol triviaal – denk aan het instellen van de thermostaat – maar vaak genoeg juist cruciaal. De recente Toeslagenaffaire is een indringend voorbeeld van wat er kan gebeuren als beslissers niet voldoende tijd nemen om stil te staan bij de gevolgen van hun keuzes. Het probleem is natuurlijk dat het lastig is om te bepalen welk besluit je wel wilt automatiseren – daarmee laat je zien dat je bereid bent om het reflectieve ‘tussenmoment’ tussen data en besluit te elimineren – en welk besluit niet. Dat vraagt volgens mij om diepgravende, kritische reflectie vanuit meerdere disciplines.

Taalgebruik

Daarmee komt ik op mijn tweede aandachtspunt: de manier waarop wij over de toepassingen van AI praten. Mijn stelling is dat ons taalgebruik ten aanzien van AI de gewenste kritische bezinning niet gemakkelijker maakt. Opnieuw een voorbeeld. Toen ik met dit artikel begon, schreef ik eerst de alinea over iemand die 's avonds na een vergadering thuis komt. Toen ik die tekst nalas, vielen mij een paar dingen in mijn eigen tekst op. Ik schreef zonder erbij na te denken dat de thermostaat ‘weet’ hoe lang de vergadering duurt en ook ‘weet’ dat het meestal wat uitloopt. En Netflix zou ‘weten’ wat zijn/haar favoriete filmgenre zo tegen elf uur 's avonds is. Even verderop maakte ik me opnieuw schuldig aan antropomorf taalgebruik. Ik betoogde dat het algoritme van Google iets ‘inschat’ en ‘tot een conclusie komt’. Letterlijk klopt dat niet, want thermostaten en Netflix ‘weten’ niet zoals mensen iets weten, en de software van Google maakt geen ‘inschattingen’ zoals mensen dat doen. Toch gebruik ik er wel die woorden voor, tenminste, als ik er niet bewust over nadenk.

Ik heb jarenlang dezelfde soort taal gehoord in het bedrijfsleven. Daar klinkt het dat ‘de data ons

zeggen dat...’ of ‘volgens het algoritme kunnen we het beste dit of dat doen’. De vraag is waar dit antropomorfe taalgebruik vandaan komt. Hebben we misschien de behoefte om een besluit of, zo u wilt, een standpunt, toe te schrijven aan iets met menselijke eigenschappen? Voelen we aan dat er ‘ergens’ een besluit genomen wordt en willen we graag ‘iets’ of ‘iemand’ hebben die we daar de verantwoordelijkheid voor kunnen laten dragen, al is het maar een algoritme? Dat laatste lukt prima in een zin als ‘*De data zeggen dat we A moeten doen.*’ Maar dan is er geen sprake meer van een bewust moment van persoonlijke verantwoordelijkheid tussen data en besluit. Heel anders klinkt het zinnetje ‘*Mijn persoonlijke conclusie op basis van de analyse van de data is dat we A moeten doen.*’ In dat geval is het immers duidelijk dat de spreker zelf degene is die verantwoordelijk is voor het standpunt.

Dit mensvormige taalgebruik bestaat overigens niet alleen in relatie tot AI. We zeggen immers net zo gemakkelijk dat de file ervoor ‘zorgt’ dat we te laat komen, of dat een rood verkeerslicht ons ‘laat stoppen’. Toch denk ik dat het goed is ons ervan bewust te zijn dat dit ook de manier is waarop we over de toepassing van AI spreken. Dat heeft te maken met het bijna alomtegenwoordige karakter ervan, en de enorme snelheid waarmee het zich blijft ontwikkelen. De impact van AI is gewoon te groot om er niet heel kritisch naar te kijken.

– Dat we data laten buikspreken, is niet iets neutraals –

Ons taalgebruik laat enerzijds zien wat onze gedachten over de toepassing van AI zijn, maar anderzijds creëert en/of bevestigt het ook bepaalde beelden. Er lijkt sprake van een concreet voorbeeld van het zogenaamde Thomas-theorema. William Thomas, een Amerikaanse socioloog, stelde in 1928 dat ‘*if men define situations as real, they are real in their consequences.*’⁸ Dus als mensen ervan overtuigd zijn dat software hun iets ‘zegt’ of ‘laat zien’, dan gaan zij vervolgens handelen alsof dat ook echt zo is.

Is dit een pleidooi voor een soort doorgeslagen *correctspeak*? Moeten we echt zo precies met onze woorden zijn? Niemand zegt immers ‘*Het stoplicht staat op rood en de betekenis die ik eraan geef is dat ik dien te stoppen.*’ Maar toch. Ons taalgebruik ten aanzien van AI kan zomaar de ruimte voor Miriam Rasch’ frictie inperken of zelfs afsluiten. Dat dienen we volgens mij niet te laten gebeuren.

U denkt inmiddels misschien dat ik anti-data of anti-automatisering ben. Dat is zeker niet het geval. Ik pleit wel voor een meer kritische houding ten opzichte van AI. Dat we door ons taalgebruik data laten buikspreken alsof data iets kunnen besluiten, vind ik niet iets neutraals. Data worden selectief verzameld en met het oog op concrete doelen geanalyseerd. De kritische vragen dienen wat mij betreft op zijn minst te zijn: Hoe werkt dat dan? Wat zijn die achterliggende doelen? En in wiens belang is die analyse? En willen we dat wel op die manier in onze samenleving?

Het gaat hier volgens mij om iets belangrijks voor ons mens-zijn. Jasmijn van der Weide schreef enige tijd geleden in de NRC⁹ dat de vraag niet is hoe menselijk we robots kunnen maken – zij bedoelt robots die zo veel mogelijk op mensen dienen te lijken – maar wat voor effect mensachtige robots op onze menselijkheid hebben. Dat geldt wat mij betreft ook voor AI. Het menselijk vermogen tot het geven van betekenis en daar verantwoordelijk voor te zijn, lijkt me een voorrecht dat we niet snel aan AI uit handen dienen te geven.

Mr. Drs. W. van der Knaap mcm is zelfstandig organisatieadviseur en docent/onderzoeker voor de Christelijke Hogeschool Ede.

Noten

1. IMEC is een onderzoekscentrum voor nano-elektronica en digitale technologie. Zie ook <https://www.imec.be/nl>.
2. G. M. De Ketelaere, *Mens versus machine*, Artificiële intelligentie ontrafeld, Pelckmans, 2020
3. M. Rasch, Frictie – *Ethiek in tijden van dataïsme*, *De bezige bij*, 2020.
4. James E. Lovelock, *Novacene. The coming of age of hyperintelligence*, Allan Lane, 2019 (Nederlandse

vertaling: *Welkom in het Novaceen – Hyperintelligentie, Gaia en de mens*, Atlas Contact, 2020).

5. J. Taylor, *Digital decisioning – using Decision Management to Deliver Business Impact from AI*, Meghan-Kiffer Press, 2019.

6. Th. H. Davenport en J. G. Harris, *Competing on analytics – the new science of winning*, Harvard Business Review Press, 2017.

7. Er bestaan veel definities van cultuur. Gangbare definities in de managementliteratuur verstaan er de gedeelde waarden, opvattingen en overtuigingen van een groep mensen onder. Een goed voorbeeld daarvan biedt de Amerikaanse auteur Edgar H. Schein.

8. Opgehaald van <https://nl.wikipedia.org/wiki/Thomas-theorema>.

9. J. van der Weide, ‘Wat doet seks met robots met de mens?’, *NRC* 10 juli 2021.