

# INHOUD

## Eerste hoofdstuk KUNSTMATIGE INTELLIGENTIE ALS PROTHESENGOD

Humanisme 7 | Machine-intelligentie 11 |  
Prothesengoden van de techniek 16 |  
Doordrongen van AI 22

## Tweede hoofdstuk KUNSTMATIGE INTELLIGENTIE VEROVERT DE TAAL

Een AI-dialoog 30 | Het einde van de  
hermeneutiek 39 | De auteur verdwijnt 42 |  
Emotioneel eenrichtingsverkeer 49

## Derde hoofdstuk LEVENDE GEEST

Voorbij het computermodel 54 |  
Emotionele en lichamelijke inbedding 55 |  
Zelfverhouding 61 | Spontaniteit 64 |

Modaliteit, temporaliteit 68 | Empathie,  
begrijpen 72 | Moraliteit 75 | Vrijheid 78

Vierde hoofdstuk  
DODE EN LEVENDE GEEST

Raadsel bewustzijn 82 | Bewuste AI? 89 |  
Nare verrassingen 92 | De carrière van de dode  
geest 96 | Afscheid van het volmaakte 102 |  
De tovenaarsleerling 107

Literatuur 111

# Eerste hoofdstuk

## KUNSTMATIGE INTELLIGENTIE ALS PROTHESENGOD

*Humanisme – Machine-intelligentie –  
Prothesengoden van de techniek –  
Doordrongen van AI*

### Humanisme

De kunstmatige intelligentie, waaraan wij ons inmiddels niet meer kunnen onttrekken, werpt haar schaduw. Het is een dynamische technologie die in staat is de menselijke geest die haar bedacht heeft radicaal in twijfel te trekken.

Sommigen verheffen deze technologie zelfs tot een religie. Dat gebeurt heel expliciet in Silicon Valley, waar een ontwikkelingsingenieur van Google onlangs verklaarde dat artificiële intelligentie (AI) weliswaar geen god is in conventionele zin, maar zich vervolgens wel afvroeg: 'Hoe moet je het anders noemen als er iets bestaat wat een miljoen keer slimmer is dan de slimste mens?' En Ray Kurzweil, een van de meest toonaangevende AI-theoretici van Silicon Valley,

verwacht dat het binnen afzienbare tijd mogelijk zal zijn om het menselijk brein met de cloud te verbinden, een ‘transformatie’ die de hersencapaciteit een ‘miljoen keer’ zou kunnen vergroten. Hier wordt een nieuwe fase van de evolutie uitgeroepen, met AI als hoofdrolspeler.

Bij dergelijke visioenen is het raadzaam om het goede oude humanisme in herinnering te brengen, dat zich ooit tegen hogere machten verzette. Nu heden ten dage de techniek voor zo’n hogere macht doorgaat waarin je moet geloven en die je moet gehoorzamen, wordt het humanistische verzet weer actueel.

Het begrip humanisme is intussen bijna een holle frase geworden, maar in de zestiende eeuw, in het tijdvak van de Renaissance, had het een scherpomlijnde betekenis. Het historische humanisme kantte zich tegen de lang in het christelijke Westen heersende opvatting dat de mens zich alleen naar God hoorde te richten, en stelde daarentegen dat de mens het potentieel heeft om zich naar zichzelf te richten: autonomie in plaats van theonomie.

Het was zaak het hogere zijn in zichzelf te zoeken in plaats van alleen in God. Het humanisme was de zelfhandhaving van de menselijke geest in grootse stijl.

Tegenwoordig zou humanisme kunnen betekenen: zelfhandhaving van de menselijke geest ten overstaan van intelligente machines. Ooit gold God als de transcendentie, tegenwoordig is het intelligente machinepark misschien een vorm van transcendentie geworden die we met ontzag en deemoed benaderen.

De werktuigen van weleer zijn denktuigen geworden waaraan wij macht over onszelf verlenen. Ten langen leste zou dan een hernieuwde humanistische opstand nodig zijn.

De beroemde Renaissancefilosoof Pico della Mirandola was zo vermetel om te beweren dat de mens een onaf wezen is, dat door God echter begiftigd is met het vermogen om zichzelf te voltooien. Wij zijn ‘wat wij willen zijn’, verklaarde hij, binnen de grenzen uiteraard van onze ‘natuurlijke beperkingen’, zoals hij eraan toevoegde. Een halffabricaat kortom, dat zichzelf kan en moet voltooien, gezamenlijk maar ook als eigen persoon.

Dit humanisme uit de zestiende eeuw was zowel maatschappelijk als individueel gericht. Je kunt zelfs zeggen: het begunstigde de onmiskenbare enkeling. De verscheidenheid van mensen gold voor het humanisme als het hoogste goed, en daarom zag het erop toe dat deze verscheidenheid ook maatschappelijk bevorderd en ontwikkeld werd. Daaraan ontsprongen normatieve ideeën die kenmerkend waren voor de moderniteit van het verlichte Europa: de vrijheid van meningsuiting en geweten bijvoorbeeld, maar ook tolerantie, rechtvaardigheid en het recht op lichamelijke integriteit.

Het humanisme verkoos een helder, positief mensbeeld. Mede daarom raakte de humanist Erasmus in een dispuut verwickeld met Maarten Luther, die de zondige natuur van de mens benadrukte en de vrije wil van de mens ontkende.

Dit positief gestemde humanisme was echter niet naïef, het erkende ook de duistere kanten van de menselijke natuur. Daarom hoort bij deze stroming ook een vindingrijkheid in het uitdenken van hulpmiddelen om de duistere aspecten van het menselijk karakter in te dammen en er zo mogelijk iets goeds van te maken. Een dergelijk hulpmiddel was bijvoorbeeld het beginsel van de scheiding der machten, dat later door Montesquieu zou worden uitgewerkt.

Het principe van de machtscheiding is beslist een van de geniaalste uitvindingen in de politiek. Het veronderstelt het inzicht dat de mens, zoals Kant zei, 'een krom stuk hout' is; daarom is een gezond wantrouwen op zijn plaats, vooral in de omgang met macht. Aangezien macht altijd misbruikt kan worden, moet zij worden opgedeeld, in principe net zoals er voor het openen van een kluis of het activeren van een wapensysteem meerdere sleutels nodig zijn die over verschillende personen en instituties zijn verdeeld. Het gaat daarbij niet alleen om een politieke scheiding tussen wetgevende, rechterlijke en uitvoerende macht. Het principe van de machtscheiding is veelomvattender en heeft betrekking op het uitbalanceren van verschillende leefsysteemen waaraan wij deelnemen en die op ons inwerken, zoals economie, godsdienst, wetenschap, techniek, politiek, openbaarheid en privésfeer. Tussen deze gebieden bestaan spanningen, die echter voor wederzijdse inperking en controle gebruikt kunnen worden. De scheiding der machten kan bescherming bieden tegen overweldiging door totale machtsaanspraken.

Tegenwoordig dreigt de techniek een dergelijke totale macht te worden. Deze situatie vraagt dus opnieuw om de humanistische wijsheid van de machtscheiding, teneinde de totalitaire aanspraak, de quasi-goddelijkheid van het intelligente machinepark, af te kunnen slaan.

## Machine-intelligentie

Bij vroegere machines werden vooral onze lichaamskrachten uitgebreid, vervangen en overtroffen. Intelligente machines dringen daarentegen door in gebieden die tot dusverre uitsluitend aan de menselijke geest werden toegeschreven: logica, cognitie, calculus, combinatie, algoritmen. Op deze gebieden wordt de menselijke geest inmiddels veruit overtroffen door zelflerende AI-systemen. We hebben te maken met prestaties waarvan we vroeger dachten dat hierbij bewustzijn moest worden verondersteld. Maar kennelijk komen ze ook zonder bewustzijn tot stand.

Als machines zonder bewustzijn intelligent kunnen zijn, verandert er iets aan het zelfbeeld van de mens.

Allereerst wijst het vermogen van intelligente machines erop dat ook wij zonder bewustzijn geestelijke prestaties leveren. We passen bij het spreken een regelsysteem toe zonder ons daarvan bewust te zijn. Niet alleen grote delen van onze lichaamsfuncties, maar ook van onze intelligentie liggen onder de bewustzijnsdrempel. Ook het identificerende kennen is

niet altijd een bewust proces. Als na verloop van tijd een bepaalde naam je weer te binnen schiet, heeft de geest blijkbaar onder de oppervlakte verder gezocht, zonder bewustzijn. Niet-bewuste intelligentie bestaat dus bij de mens en sinds kort ook in machines.

Halverwege de vorige eeuw probeerde Norbert Wiener met zijn cybernetica de niet-bewuste intelligentie te vatten. Dat geest en machine op vergelijkbare wijze functioneren, komt volgens deze theorie door het terugkoppelingsprincipe. Dit houdt in, aldus Wiener, 'dat het gedrag aan de uitkomst wordt getoetst, en dat het intreden dan wel uitblijven van die uitkomst het toekomstige gedrag beïnvloedt'. Zulke terugkoppeling (of feedback) voltrekt zich bij de mens aanvankelijk intuïtief (in de kindertijd), wordt later bewust aangeleerd en kan uiteindelijk zonder bewuste sturing routinematig en automatisch verlopen. Dit proces kan worden voorgesteld als een informatieverwerkingsproces en worden nagebouwd in een machine met een complex feedbacksysteem van sensoren en besturingselementen. Het functioneren van dergelijke apparaten wekt de indruk van intelligentie.

Dat geldt des te meer nu de digitalisering in volle gang is, waarbij informatie uit schriftelijke en andere analoge beeld- en geluidsmedia in digitale code wordt omgezet en voor elektronische dataverwerking beschikbaar wordt gemaakt. Zo bestaan er hele waarnemings- en gedachtewerelden in digitale representatie, dat wil zeggen in een uniforme universele taal van nullen en enen, gevormd in de elektronische

schakelvolgorde van aan en uit.

Dit tekensysteem maakt het mogelijk om volgens vastgelegde algoritmen logische operaties, calculaties, combinaties en berekeningen uit te voeren met de in digitale tekens getransformeerde kennis van de wereld. Dat resulteert in een enorm potentieel voor steeds nieuwe koppelingen van informatie in machinaal leesbare patronen. De feedbackprocessen zorgen daarbij voor een exponentiële groeidynamiek.

Zo ontstaat een omvangrijke opslag- en communicatiewereld, een tweede werkelijkheid die uit de direct zintuiglijk waarneembare eerste werkelijkheid voortkomt, daar met de traditionele geestelijke wereld van de mens in aanraking komt en zich ermee vermengt.

Deze nieuwe technologische werelden bouwen eveneens voort op menselijke vermogens die versterkt of vervangen worden, zoals bij een kraan of een verkeersmiddel. In dit geval zijn het echter de geestelijke vermogens die ondersteund, versterkt of vervangen kunnen worden: het vermogen tot rekenkundig denken, onthouden, combineren, redeneren, calculeren, het herkennen van patronen. Met dat alles lijkt de carrière van de machine-intelligentie nog maar net begonnen te zijn.

De ervaring met machinale intelligentie scherpt de blik voor de gebieden van de menselijke intelligentie waar het er in zekere zin altijd al machineachtig aan toeging, in tegenstelling tot bijvoorbeeld het spontane, willekeurige, fantasierijke, emotionele. Het zijn de gebieden van strikte rationaliteit – de logica, het

rekenen, de accuraat gereguleerde stappenreeks van algoritmen, vaardigheden bovendien die je ook nodig hebt voor de bouw van machines.

In de zelfervaring van de geest gaat deze strenge rationaliteit meestal met het gevoel gepaard dat je haar niet spontaan volgt maar je er met moeite aan onderwerpt. Op de paden van de logica kan het denken nu eenmaal niet vrij bewegen zoals het wil. Een dwang is voelbaar. De indruk van het vreemde, het levenloze bedriegt niet. Want inderdaad lijkt het denken in de sporen van de logica op een machineachtig proces, dat vervolgens zijn externe weergave vindt in het functioneren van de machine.

Geconfronteerd met de eerste computers in de jaren vijftig ontwikkelde Gotthard Günther, een filosofische pionier op het gebied van kunstmatige intelligentie, het idee van een driewaardige ontologie: ten eerste organisch leven, ten tweede anorganische dingen en sinds kort ten derde, in de vorm van cybernetische machines en computers, intelligente dingen. Dienovereenkomstig zou de – in het organische gewortelde – levende geest van de mens concurrentie hebben gekregen van de dode geest van de intelligente machine.

Freud sprak ooit van de drie krenkingen van het menselijk narcisme. De eerste krenking: Copernicus – de mens kan zichzelf niet langer in het middelpunt van het heelal lokaliseren. De tweede krenking: Darwin – de mens stamt af van de aap. De derde krenking: de psychologie van Freud zelf – zijn ontdekking van