



Rechtspreken lukt een computer nog niet echt.

Oude deugden actueel voor nieuwe technische wereld

## ‘TECHNOLOGIE IS NIET IETS WAT ONS OVERKOMT’

We worden omgeven door intelligente machines in allerlei soorten en maten. De opkomst van slimme machines stelt ons voor grote ethische vraagstukken en grijpt diep in op ons mens-zijn. Hoe kunnen menselijke waarden centraal komen te staan in de ontwikkeling van kunstmatige intelligentie?

Tekst: Elleke Bal

**A**dvocaten en rechters keken reikhalzend uit naar een experiment tijdens de jaarlijkse Dag van de Rechtspraak in 2017. Voor heel even kreeg een computer een rechtershamer in handen. Een computersysteem gemaakt door een bedrijf gespecialiseerd in juridische da-

ta-analyse kreeg twee zaken voorgelegd. Als eerste een arbeidszaak, waarin een financieel directeur zijn ontslag aanvecht en een ontslagvergoeding eist. De tweede zaak draaide om een voorval waarbij een automobilist een rotonde niet rechtsom maar linksom neemt en daarvoor een boete krijgt. Om een gok te doen over het vonnis

doorzocht de computer met behulp van bepaalde algoritmes data van vergelijkbare zaken uit het verleden.

In de verkeerszaak zat de computer eraan. Die dacht dat de man schuldig was aan de verkeersovertreding, maar miste een cruciaal detail. Het verkeersbord bij de rotonde ontbrak, waardoor de automobilist

onterecht de boete had gekregen, oordeelde een rechter van vlees en bloed. In de arbeidszaak deed de computer het iets beter. Net als een menselijke rechter oordeelde ook de computer dat de man onterecht was ontslagen. Maar de ontslagvergoeding die de menselijke rechter toekende was € 65.000, terwijl de computer daar € 6.500 van had gemaakt. “Al met al vind ik dat de computer nog lang niet kan tippen aan het ambacht van de rechter”, sprak rechter Erik Koster, die het in dit experiment opnam tegen de computer.

Volgens de Raad van de Rechtspraak liet het experiment zien dat het ambacht van de rechter noodzakelijk is. Machines kunnen nog geen zelfstandige besluiten nemen, laat staan dat ze een vonnis motiveren. Maar toch worden wereldwijd stukje bij beetje onderdelen van het werk van rechters en advocaten geautomatiseerd. Er zijn al zelflerende programma's die grote hoeveelheden data doorspitten om zo advocaten te helpen in de voorbereiding van zaken en er zijn systemen die een percentage kunnen produceren dat de kans



**'We moeten proberen heel goed te begrijpen hoe techniek de maatschappij en de mens verandert'**

op recidive weergeeft. Het zal niet lang meer duren voordat computerprogramma's argumenten voor en tegen een veroordeling van een verdachte op een rij kunnen zetten. Maar wie heeft inzage in deze gegevens? En zijn deze computersystemen wel gestoeld op menselijke waarden zoals gelijkheid?

**Ethiek**

De rechtspraak is slechts een van de vakgebieden die de komende jaren ingrijpend geraakt zal worden door de toepassing van artificiële intelligentie (AI). Dat gegeven brengt grote ethische vraagstukken met

zich mee. Volgens de Europese Commissie moeten 'Europese waarden' centraal komen te staan in de ontwikkeling van zinvolle en betrouwbare kunstmatige intelligentie in Europa. In april van dit jaar kwam een speciale Europese deskundigengroep van 52 experts met een opvallend en uitgesproken advies. Zorg voor 'ethische AI', schreven zij. De Europese experts munten ook de term *human-centric AI*, vertelt Roos de Jong, techniekfilosoof en onderzoeker bij het Rathenau Instituut. Dat betekent zoveel als: stel de mens centraal. De Jong: "AI zou geen doel op zich moeten worden, maar we moeten het inzetten om het menselijk welzijn te vergroten. De Europese waarden vormen daarvoor de kern." Een fundamenteel andere benadering dan in China, merkt ze op, waar de overheid een monopolie heeft op kunstmatige intelligentie en technologieën veel worden gebruikt om burgers in de gaten te houden. Of neem Amerika, waar de grote techbedrijven deze technologieën gekaapt hebben om consumenten aan zich te binden. Volgens De Jong zou het proble-



Een robot kan ook persoonlijke aandacht geven.

matisch zijn om de grote bedrijven achter kunstmatige intelligentie ook hier te veel ruimte te geven. Juist in een democratische rechtsstaat moeten maatschappelijk debat en wetgeving een centrale rol spelen bij de inzet van AI. Het gaat om de vraag hoe we onze samenleving willen organiseren? Dat AI op een fundamentele manier in ons leven kan ingrijpen wordt steeds breder erkend. Daarom zijn er binnen de wetenschap, het bedrijfsleven, de overheid en maatschappelijke organisaties veel initiatieven om vanuit ethisch perspectief naar AI te kijken.

Zo werkt Marc Steen bij onderzoeksorganisatie TNO aan de ethische vraagstukken die komen kijken bij technologische innovatie. Het is zijn taak om samen met collega's kritische vragen te stellen over de maatschappelijke of ethische implicaties die een bepaalde uitvinding zou hebben op de levens van mensen. "Stel dat een collega mij vertelt over een bepaald nieuw product of een computerprogramma. Dan stel ik kritische vragen zoals: zijn er ook publieke waarden in het geding? Hoe zit het met privacy? En wat als dit idee zich over de wereld zou verspreiden, hoe zou de wereld er dan uit zien?" Volgens Steen zal het de komende jaren steeds gewoner worden om dit soort vragen te stellen, met name als het om kunstmatige intelligentie gaat.



Roos de Jong:  
"Kunstmatige intelligentie moet geen doel op zichzelf worden."

De filosofisch-ethische traditie vindt hij daarin waardevol.

Dat beaamt Roos de Jong. "Ik moest vroeger vaak verdedigen waarom ik filosofie ging studeren, maar dat is nu niet meer nodig", lacht ze. Dat komt volgens haar deels omdat steeds meer mensen zich realiseren dat AI diep ingrijpt op ons menszijn. "We worden omgeven door intelligente machines in allerlei soorten en maten. We worden gemonitord en gestuurd. Wat betekent het wanneer we op onbewust niveau worden beïnvloed en keuzes voor ons worden gemaakt door onze slimme omgeving? Zijn we dan nog in staat om onze eigen doelen en waarden te bepalen en na te streven?"

### Beperkte intelligentie

Overigens bestaan er volgens De Jong veel misverstanden over wat AI nu precies is. Je kunt het zien als de motor van het tweede machinetijdperk, zegt ze. Alleen leveren deze machines geen spierkracht zoals tijdens de industriële revolutie, maar denkkracht. Zo breed als het begrip 'intelligentie' is, is ook het concept kunstmatige intelligentie. De test die wel eens gebruikt wordt om vast te stellen of een computer intelligent is, is de Turingtest. Als een computer een mens voor de gek kan houden en 'm kan laten denken dat hij een mens is, móét er wel sprake zijn van intelligentie.

In sciencefiction zie je vaak systemen die de menselijke intelligentie overstijgen en vervolgens de hele mensheid overnemen, vertelt De Jong. "Maar de meeste systemen kunnen maar op een heel beperkt vlak intelligent gedrag vertonen." Ze noemt het voorbeeld van de computer AlphaGo van Google, die heel goed is in het Aziatische spel Go. Computervariant AlphaZero blijkt van elke schaakmeester te winnen. Maar voetballen, tennissen en monopoly spelen kunnen computers niet. Daarnaast is AI niet één technologie, maar vaak een onderdeel van een cluster van technologieën zoals sensoren, *big data* en algoritmes die ervoor zorgen dat machines in een bepaalde omgeving bepaald gedrag kunnen vertonen. Toch is AI niet per se futuristisch; de meesten van ons worden er al dagelijks door beïnvloed. Spamfilters in onze mailboxen werken op

basis van kunstmatige intelligentie, net als de aanbevelingen die je krijgt als je Netflix kijkt, Spotify luistert of iets koopt in een internetwinkel. Onze voorkeuren worden constant bijgehouden en gemonitord. En razendsnel wordt deze informatie weer gebruikt om ons nog beter 'van dienst te zijn'. "AI wordt vaak gekenmerkt door de snelheid van informatieverwerking", vertelt De Jong. Ze begrijpt dat daardoor soms het idee kan ontstaan dat zo'n machine het menselijk kunnen te boven gaat en dat we aan alle kanten ingehaald en gestuurd worden. Dat was ook waar Stephen Hawking in 2014 voor waarschuwde in een interview met de BBC. "De ontwikkeling van kunstmatige intelligentie kan het einde van de mensheid betekenen", zei hij. "De techniek verbetert zichzelf op een steeds hoger tempo. De mens kan die groei niet bijhouden, omdat biologische evolutie hen tegenhoudt."

De Jong is er niet zo bang voor. "Mensen zijn meer dan 'informatieverwerkers'. Een zesjarig kind kan nu in veel opzichten nog veel meer dan een computer." Maar, zegt ze ook: we hebben nu wel van alles te beslissen, dus het is nu belangrijk dat we als samenleving nadenken over de vraag: in wat voor soort wereld willen we leven? "Technologie is niet iets wat ons overkomt", zegt ze, we bouwen haar zelf, en zij kan ons helpen bij het bereiken van onze doelen. Maar aan de andere kant beïnvloedt diezelfde technologie hoe we naar de wereld kijken. "We moeten dus proberen heel goed te begrijpen hoe techniek de maatschappij en de mens verandert. Ik denk dat we juist daardoor verantwoorde technologie kunnen ontwikkelen."

Wie eerst bepaalt welke waarden centraal staan, kan die waarden vervolgens meenemen in de ontwikkeling van bijvoorbeeld robots, zegt De Jong ook. Dan kun je in de zorg bijvoorbeeld zeggen, dat je 'persoonlijke aandacht' belangrijk vindt. "Dan blijf je ook niet hangen in de discussie: meer of minder technologie. Ik geloof dat we veel meer moeten denken in termen van samenwerking." Dat betekent volgens De Jong wel dat het hardnekkige onderscheid in alfa- en bèta-opleidingen moet worden opgeheven. "We hebben bijvoorbeeld ingenieurs nodig die ook over waarden, ethiek, filosofie en mens-zijn nadenken."



Marc Steen: "Zijn er ook publieke waarden in het geding? Hoe zit het met privacy?"

### Deugden

Marc Steen van TNO verwijst graag naar het werk van een Amerikaanse filosofe die zich heeft gespecialiseerd in technologie-ethiek: Shannon Vallor. Zij schreef het boek *Technology and the Virtues* (Technologie en de deugden), waarin ze bepleit dat het tijd is om ons weg te bewegen van twee extremen: AI gaat óf al onze problemen oplossen óf het brengt de mensheid om zeep. "Het wordt iets ertussenin. Het hangt ervan af welke deugden wij cultiveren en manifesteren tijdens het ontwikkelen en toepassen van AI."

Net als Vallor betoogt Steen – opgeleid als industrieel ontwerper in Delft – dat we deugdethiek hard nodig hebben bij vragen over opkomende technologieën. Volgens Aristoteles is rechtvaardigheid de belangrijkste deugd, zegt Steen. De deugdethiek gaat over het inrichten van de maatschappij op zo'n manier dat mensen kunnen floreren. Deze denkwijze is een van de ethische tradities naast plichtethiek en gevolgenethiek. Die laatste twee zijn alleen minder geschikt, vindt Steen, omdat plichten en regels rondom AI nog helemaal niet duidelijk zijn. "AI is nog volop in ontwikkeling. Maar onze deugden zijn juist vrij stabiel."

Zo vindt hij de deugd bescheidenheid een belangrijke. "Heel veel technologieën staan



### 'Neem de waarheid van het algoritme niet als een absolute waarheid'

nog in de kinderschoenen, dus daar moet je je bewust van zijn als je zo'n technologie wilt toepassen of gebruiken." Rechtvaardigheid is een andere deugd die volgens hem van pas kan komen: "Stel je werkt in een bedrijf en je maakt een bepaald algoritme: kun je straks thuis uitleggen: wat heb ik vandaag gedaan? Heeft het iets bijgedragen?"

Steen houdt zich dagelijks bezig met de ethische vragen die komen kijken bij technologische innovatie. Zo deed hij onlangs een week lang een project met de politie. Ze spraken over een aantal potentiële innovaties zoals een computerprogramma dat een risicotaxatie kan maken voor 'herhaald slachtofferschap'. Hoe groot is de kans dat een slachtoffer nog eens getroffen wordt? Die vraag is nog weinig gesteld omdat tot nu toe de nadruk vooral op herhaald daderschap heeft gelegen, maar is minstens zo interessant. Zo'n beoordeling zou de politie kunnen helpen om kwetsbaarheid vast te stellen en iemand verdere hulp te bieden. Maar het vraagt wel om ander soort registraties in politiestystemen.

"Daar komen toch wel wat ethische vragen

bij kijken", zegt Steen. "Het gaat dan om hoe je iemand benadert. Stel, iemand komt aan de balie, en een computer gaat kijken naar de mogelijkheid dat deze persoon binnenkort wéér slachtoffer wordt. Dan is bescheidenheid belangrijk. Je weet bijvoorbeeld dat er altijd valse positieven uit een algoritme kunnen komen, en daar kun je rekening mee houden. Loop in twijfelgevallen niet te hard van stapel. Neem de waarheid van het algoritme niet als een absolute waarheid."

Om niet zomaar klakkeloos een nieuw product, een nieuwe machine of een nieuw algoritme op de samenleving los te laten, dat vraagt ook moed, zegt Steen – nog zo'n deugd. "Moed om aan de rem te trekken, tijd te nemen om iets beter na te denken over het inzetten van AI." Dat zou voor bedrijven die zich vooral bezighouden met spullen verkopen een mooie kunnen zijn. Is het niet een beetje een ideale wereld die Steen omschrijft? Want tot nu toe lijkt het alsof de techgiganten als Facebook, Twitter en Amazon zich toch vooral druk maken hun omzet, in plaats van over deugden. "Dat klopt wel", zegt Steen. "Het is ook vooral een hoop die ik heb." Maar die hoop wordt bij hem zeker gevoed door de nieuwe visie van Europa op het ontwikkelen van *human-centric AI*. "Dat is voorlopig een van de beste pogingen wereldwijd om hier iets mee te doen. AI zou een middel moeten zijn voor het menselijk floreren en het waarborgen van mensenrechten en menselijke waardigheid." ●



Computer AlphaZero wint elke schaakwedstrijd.