

‘We hebben geen slimme machines nodig, maar wijze’

Deze publicatie is onderdeel van het thema [Over taal gesproken](#) op Kennislink.nl.

Wetenschapsfilmfestival InScience stond afgelopen weekend in het teken van kunstmatige intelligentie. De kersverse hoogleraar Marcel van Gerven verkent hoe onderzoek naar kunstmatige en natuurlijke intelligentie elkaar kan inspireren en versterken.

Auteur: [Erica Renckens](#)

In de film *Her* (2013) valt hoofdpersoon Theodore voor de charmes van de virtuele assistente in zijn telefoon. Hij wordt hopeloos verliefd en weet zich geen raad als blijkt dat het intelligente besturingssysteem in staat is met duizenden mensen tegelijk relaties te onderhouden. Samantha, de virtuele assistent met de zwoele stem van Scarlett Johansson, leert van alle input die ze krijgt en ontstijgt daardoor al snel de grenzen van het menselijke brein.



Trailer van de film *Her* uit 2013, die tijdens wetenschapsfilmfestival InScience werd vertoond.

De film, die afgelopen weekend werd getoond tijdens wetenschapsfilmfestival [InScience](#), toont een doemscenario waarin kunstmatige intelligentie ontspooit. Heel realistisch is dit gevaar echter nog niet, zo vertelde de kersverse hoogleraar Kunstmatige intelligentie Marcel van Gerven begin september tijdens zijn inaugurele rede: “Menselijke machines zijn dan wel een theoretische mogelijkheid, maar praktisch gezien zijn we hier nog ver van verwijderd. Bovendien is het nog maar de vraag of zulke machines dezelfde subjectieve ervaringen zullen hebben als mensen.”

Een veel urgenter gevaar is volgens hem onze overschatting van de mogelijkheden van kunstmatige intelligentie. “Huidige, nog relatief domme intelligente technologie grijpt al regelmatig in de hedendaagse samenleving in. Ze worden ingezet om zowel onze waarneming als gedrag te monitoren, te manipuleren of zelfs volledig over te nemen.” Als voorbeelden noemt hij de opkomst van apps die niet van echt te onderscheiden nepnieuws kunnen creëren, het



Marcel van Gerven: “Slimme systemen zijn eigenlijk totaal niet slim.” *Radboud Universiteit*

[sociaal kredietsysteem](#) dat China wil invoeren waarbij elke burger een score krijgt op basis van zijn gedrag, en autonome wapens die zelf kunnen beslissen over leven en dood.

RU, MPI

Complexe afwegingen

“Zulke systemen zijn eigenlijk totaal niet slim en niet in staat om na te denken”, vertelt Van Gerven in zijn werkkamer. “Dat komt doordat de intentionaliteit ontbreekt, de gedachte achter een daad. Volgens mij staan we aan de start van een nieuwe revolutie waarbij psychologen en neurowetenschappers moeten samenwerken met informatici en fysici om de principes die ons brein kenmerken, na te bootsen in machines.” Zo kunnen machines wel over de intentionaliteit beschikken die mensen kennen. “We moeten machines een wil geven, een drive, waardoor ze interne modellen ontwikkelen van zichzelf en hun omgeving waarmee ze kunnen nadenken.”

Machines met een eigen wil? Moeten we dat echt willen? “De vraag is natuurlijk hoe ver je het wilt doorvoeren. Misschien willen we eigenlijk helemaal geen slimme machines bouwen, maar wijze machines. Machines die kwaliteiten hebben als empathie en compassie.” Om dat te bereiken zouden volgens Van Gerven naast gamma- en bèta-wetenschappers ook alfa’s bij de ontwikkeling betrokken moeten zijn.

De systemen die dat zou opleveren, zijn volgens Van Gerven zeker nuttig. “Elke technologie heeft positieve en negatieve kanten. Op het moment dat we zelfrijdende auto’s kunnen maken die minder ongelukken veroorzaken dan de mens, is het de vraag wat je wil: een auto die zelf beslissingen neemt en af en toe iemand aanrijdt, of niet – en dan onder de streep meer verkeersdoden. Dat zijn complexe afwegingen, waar je je ook bewust van moet zijn als je zulke systemen bouwt. Gelukkig hebben we in onze vakgroep [Pim Haselager](#) om dat bewustzijn te verankeren.” Haselager is naast expert op het gebied van kunstmatige intelligentie ook filosoof, en fungeert zo als een soort moreel geweten van de vakgroep.

Synthetisch brein

Maar zo ver is het allemaal dus nog niet. Waar houdt Van Gerven zich dan wel zoal mee bezig? “Mijn eerste interesse is het begrijpen van de menselijke geest. Dat wil ik doen door exacte, wiskundige technieken uit de kunstmatige intelligentie toe te passen”, vertelt hij. Aan de hand van intelligente machines probeert hij te ontrafelen hoe het brein werkt. “Het is natuurlijk leuk als je kunt laten zien welk hersengebied actief wordt bij een bepaalde taak, maar dat vertelt ons helemaal niets over hoe het werkt. Wij bouwen modellen die ons dat wél vertellen. Die slimmere technieken stellen ons ook in staat om gedachten uit te lezen en daarop in te grijpen.”

Daarom werkt hij met zijn collega’s aan brein-geïnspireerde computermodellen, die werken als echte menselijke hersenen. “We bouwen modellen die zich hetzelfde gedragen als wijzelf. Daarmee voeren we experimenten uit en zo krijgen we meer inzicht in ons eigen gedrag. We willen zo leren wat er gebeurt vanaf het moment dat informatie het brein binnenkomt tot het is verwerkt: hoe worden die neurale representaties door het brein gesluist? Dat willen we onderzoeken voor wat je ziet en voor taal die je verwerkt en produceert.”

Aan een roze olifant denken

Op het gebied van visuele waarneming hebben Van Gerven en zijn collega’s al grote stappen gezet. Ze ontwikkelden machine learning-technieken die op basis van gemeten hersenactiviteit reconstrueren wat iemand ziet. “Dat is eigenlijk een eerste stap”, aldus Van Gerven. “Ons systeem reconstrueert het beeld dat

iemand op zijn netvlies heeft, maar uiteindelijk willen we interne representaties kunnen uitlezen, zoals datgene wat iemand zich inbeeldt.” Dit biedt niet alleen inzicht in verschillende hersenprocessen, maar maakt het op termijn ook mogelijk om te communiceren op basis van onze gedachten. “Je hebt een combinatie van intelligente technologie en betere neurotechnologie nodig om hersenactiviteit steeds gedetailleerder uit te lezen.”

Waarneming



Van Gerven en zijn collega's ontwikkelden computer-modellen die op basis van hersenactiviteit reconstrueren wat we hebben waargenomen. De bovenste rij zijn de werkelijke gezichten, de onderste de reconstructies van de computermodellen.
Van Gerven

Reconstructie

Als je die hersenactiviteit kunt uitlezen, is de volgende stap haar manipuleren. “We kunnen specifieke neuronen in het visuele veld stimuleren, wat zorgt voor lichtflitsen. Daarmee kun je patronen genereren en uiteindelijk blinden een deel van hun waarneming teruggeven. Dan gaat het om eenvoudige beelden, zoals de omtrek van een gezicht.” Momenteel onderzoekt Van Gerven wat de optimale stimulatiepatronen zijn om ziende mensen te laten zien wat een camera registreert.

De wetenschappers hebben contact met Spaanse neurochirurgen die waarschijnlijk binnen afzienbare tijd voor het eerst bij mensen elektroden in de hersenen zullen implanteren. “Zij krijgen een grid met duizend elektroden op hun primaire visuele schors van het brein. Elke elektrode wekt ergens in het zichtveld een lichtflits op”, vertelt Van Gerven. Dankzij ontwikkelingen in de kunstmatige intelligentie wordt het resultaat daarvan steeds beter, voorspelt hij. “Stel, ik wil het contextafhankelijk maken: als ik met jou spreek, wil ik je gezichtsuitdrukking zien en niet de hele omgeving. Dan moet het systeem daarvoor de relevante informatie detecteren.”

Interne spraak uitlezen

Van Gerven is ook betrokken bij het onderzoeksproject [Language in Interaction](#), dat het menselijk taalvermogen probeert te doorgronden. “Iets vergelijkbaars wat we nu al doen met zicht, willen we daar doen met taal. We willen decoderen en representeren welke semantische representaties je in je brein hebt. Dan kunnen we ook interne spraak uitlezen.”

Komt dat niet angstvallig dicht bij gedachten lezen? “Op dit moment formuleer ik zinnen in mijn hoofd; die zal je kunnen uitlezen. Niet mijn kennis, maar alleen wat ik zelf actief maak in mijn brein. Maar het is zeker belangrijk om na te denken waarom we dit doen. Wetenschappelijk gezien is het interessant en levert het ons veel nieuwe inzichten op. En je laat er verlamde mensen weer mee communiceren. Maar er schuilt natuurlijk ook het gevaar van misbruik in, dat moet je met bijvoorbeeld wetgeving beheersen.”

Dromen van elektrische schapen

Om deze techniek mogelijk te maken heeft Van Gerven eerst nog grote hoeveelheden data nodig: gemeten hersenactiviteit van mensen van wie bekend is wat ze op dat moment horen, bijvoorbeeld enkele uren aan luisterboeken. “Voor wat we zien hebben we die data wel al, daarom ligt daar nu ook nog meer de focus op. Daar hebben we bijvoorbeeld ook zelf data verzameld: een masterstudent heeft dertig uur in een hersenscanner naar afleveringen van Dr. Who gekeken.”

Maar dan heb je toch maar de neurale responsen van één persoon? “Generaliseren is belangrijk om algemene uitspreken te kunnen doen over een populatie. Maar als je gemiddeldes neemt over meerdere breinen, zijn de activatiegebieden die je nog ziet slechts het topje van de ijsberg. Ik begrijp liever één brein in die mate dat ik het kan nabouwen, dan dat ik alleen gemiddeldes over alle breinen kan observeren.”

Het werk van Van Gerven zet kleine stapjes op de weg richting menselijke machines. Maar dat eindpunt is nog lang niet bereikt, zo sloot Van Gerven zijn oratie af. “Er zijn nog steeds geen androïden die dromen van elektrische schapen”, zei hij, verwijzend naar de [science fiction-roman](#) uit 1969. De mens zelf is echter voor Van Gerven het grote wonder: “Hoe de geest kan ontstaan uit levenloze materie is voor mij de ultieme wetenschappelijke vraag.” Dat maakt het menselijk brein nog altijd het grootst te ontrafelen mysterie.



Om specifieke hersenactiviteit te kunnen koppelen aan de visuele waarneming van bepaalde beelden, keek een student dertig uur in een fMRI-scanner naar afleveringen van de serie Dr. Who.

Barbara Braams