

# De **quiz** die al decennialang ons denken over **AI** stuurt

**Mens vs computer** | Moet je bewustzijn hebben om intelligentie te vertonen? De Turingtest is bedoeld om vat te krijgen op die vraag. Nu we worden overspoeld met de maaksels van grote taalmodellen, lijkt die relevanter dan ooit.

tekst **Joost van Egmond**

**W**e maken er een quiz van. Aan de ene kant van het gordijn zit de deelnemer, aan de andere kant een mens en een computer. De deelnemer stelt vragen aan beide en moet op basis van de antwoorden raden wie wie is. De taak van de computer is om het de deelnemer zo moeilijk mogelijk te maken. Kan een computer dat even goed als een mens?

De beroemde wiskundige Alan Turing maakte het destijds iets ingewikkelder, maar in wezen is dit de kern van de Turingtest, een criterium om antwoord te geven op de vraag: kunnen machines denken?

Sinds het in 1950 als een toegankelijk artikel verscheen in het wetenschappelijke tijdschrift *Mind*, is de Turingtest een onlosmakelijk deel geworden van zowel de *nerdy* populaire cultuur als van veel wetenschappelijke verhandelingen over kunstmatige intelligentie en bewustzijn.

Het achterliggende idee: we kunnen lang debatteren over wat denken is, maar uiteindelijk beoordelen we een proces aan de hand van de uitkomst. En als die uitkomst van een computer niet meer is te onderscheiden van die van een mens, zul je moeten concluderen dat de computer denkt.

Dat roept nogal wat weerstand op. Een vaak terugkerend probleem is of je al dan niet onderscheid kunt maken tussen denken en bewustzijn. Bij mensen gaat dat uiteraard hand in hand. Wij voelen onszelf denken, zoals Descartes voor het haardvuur zat en concludeerde: 'Ik denk, dus ik ben'. Wij kunnen ons weinig voorstellen bij denken zonder de bijbehorende ervaring dat je denkt. Maar is die ervaring daarmee

ook noodzakelijk voor denken?

Die vraag schuift de Turingtest terzijde.

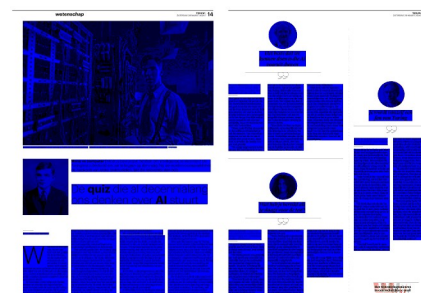
Turing wees er zelf al op dat de enige manier om zeker te zijn dat een machine daadwerkelijk ervaart dat die denkt, is om die machine zelf te zijn, en zelf dat denkproces te ervaren. Wij mensen kunnen het ervaren van bewustzijn ook alleen bij onszelf vaststellen, niet bij een ander. Maar om niet eeuwig te bakkeleien of andere mensen wel denken, argumenteerde Turing, is het gebruikelijk om ervan uit te gaan dat ieder mens denkt. Eenzelfde pragmatische keuze stelde hij voor bij het beoordelen van een computer.

Het is niet zo dat programmeurs de hele dag zitten te sleutelen om taalmodellen te optimaliseren voor de Turingtest, daar zijn heel andere criteria voor. Sowieso zijn de regels te vaag. Hoe selecteer je de deelnemer en hoeveel moeten er de mist ingaan voor de computer heeft gewonnen?

In letterlijke zin is de Turingtest dan ook meer iets voor feesten en partijen, maar als concept heeft dit gedachtenexperiment een gigantische invloed gehad op ons denken over denken, en heeft het hele generaties ontwikkelaars gemotiveerd.

Maar gaat dat concept nog op? Turing was een echte pionier. Hij ontwikkelde deze argumentatie in een tijd dat ook de beste computers nog maar een fractie van de huidige rekenkracht hadden. Maar een vrijblijvende fantasie was het zeker niet. De verwachting dat computers snel beter zouden worden in dit spel was er toen ook al. Turing zelf waagde zich aan de voorspelling dat rond de eeuwwisseling een gemiddelde deelnemer nog maar 70 procent kans zou hebben om het spel binnen vijf minuten winnend af te ronden.

Dat is niet bewaarheid. Is het een kwestie van meer tijd, of is er een fundamenteel probleem met de redenering? En hoe zinnig is het driekwart eeuw later nog om langs deze lijnen te denken? *Trouw* vroeg een aantal uiteenlopende computerwetenschappers of de Turingtest nog nut heeft.





## *Het beste dat we kunnen doen is die AI vooruit duwen*

**Thomas Moerland** AI-onderzoeker aan het informatica-instituut Liacs en auteur van het boek 'Van IQ naar AI'

“Het punt is niet dat we intelligentie en bewustzijn per se als iets verschillend moeten zien, het gaat erom dat we maar één van de twee kunnen meten. Er is geen manier om vast te stellen of mijn telefoon iets meemaakt als ik iets intyp. Het lijkt mij in eerste instantie ook onwaarschijnlijk, maar we kunnen het niet echt vaststellen.

“Maar zoals Turing al aanstippte: datzelfde probleem heb je ook bij mensen. Van onszelf accepteren we het als we dat eigenlijk niet goed weten, maar over AI worden die lastige vragen wel gesteld.

“De enige weg vooruit is dus om aan die intelligentie te wer-

ken en te zien hoever we daarmee komen. Dat is het geniale van Turing. Hij stelde voor om niet te blijven discussiëren over wat die subjectieve ervaring is, maar te kijken naar wat de computer kán. We kijken dan dus alleen naar wat we kunnen meten. We zijn bezig om die computer iets slims te laten doen, met bewustzijn houden we ons niet bezig. Dat betekent ook dat we mogelijk dingen missen, maar het is wel de manier waarop je vooruit komt.

“Hoe dicht we dan nu bij intelligentie op menselijk niveau zijn is lastig in te schatten. We zijn al vaak te optimistisch geweest. We hebben moeite om in te schatten hoe ingewikkeld we zelf in elkaar zitten. Het zou best kunnen zijn dat het ‘recept’ van de huidige generatie AI niet voldoende is en dat er iets heel

nieuws moet bijkomen.

“Het beste dat we kunnen doen is die AI vooruit duwen en kijken hoever we komen. Bij dat onderzoek speelt de Turingtest nauwelijks een rol. Ik ken geen onderzoeker die de hele tijd bezig is om computers aan de Turingtest te onderwerpen. De grootste praktische toepassing ervan zijn eigenlijk die opdrachten die je online soms krijgt om te bewijzen dat je geen robot bent. Daarmee doen we allemaal de omgekeerde Turingtest: kan ik een handeling tonen waar een computer nog moeite mee heeft? Je ziet dat die opdrachten in de loop der tijd steeds ingewikkelder worden. Daar zit de vooruitgang.”



## *Wat heb je bereikt als je slaagt voor de test?*

**Malvina Nissim** hoogleraar computationele taalkunde en maatschappij aan de Rijksuniversiteit Groningen

“De grote vraag is wat je nu eigenlijk aan het testen bent. Stel, je slaagt voor de test, wat heb je dan bereikt?”

“De Turingtest is een gesprek, dus taal speelt een centrale rol. Daar zijn met grote taalmodellen nu enorme stappen in gemaakt. Maar het is belangrijk om voor ogen te houden wat die modellen doen. Ze krijgen taalinput, ze maken berekeningen en genereren taaloutput.”

“Nou ben ik gek op taal, maar we kunnen dat niet volledig gelijkstellen aan menselijke intelligentie. Taal is een heel belangrijke en krachtige manier waarop we ons denken uit te

drukken, maar denken is niet beperkt tot taal.

“Toen Turing dit experiment bedacht, was het een droom dat een machine vloeiend zou kunnen zijn in taalgebruik. Voor de huidige grote taalmodellen is vloeiend taalgebruik niet langer het probleem. Maar desondanks is het nog steeds niet moeilijk om het te onderscheiden van menselijk taalgebruik.

“Je kunt proberen om daar verdere vooruitgang in te boeken. Dat zal ons dingen leren over taal, en hoe we taal gebruiken. We kunnen zo begrijpen hoe een model taal verwerft en wat het met die taal kan doen. We zien bijvoorbeeld dat modellen abstracties leren uit taal, en we weten nog niet goed hoe ze dat doen. Als onderzoeker ben ik daarin geïnteresseerd, maar ik

denk niet dat dit de weg is die we moeten begaan om intelligentie te begrijpen. Als we meer over onszelf en ons bewustzijn willen leren, zou ik eerder inzetten op neurowetenschap en cognitiewetenschap. Ook blijft het belangrijk om taal in echte menselijke interactie te bestuderen.

“Dus de vraag blijft: wat wil je dat zo'n machine doet? Waarom zou je een machine willen maken die in een gesprek perfect een mens na kan doen? Als je het mij vraagt, wil je juist dat die machine duidelijk van een mens te onderscheiden is.”



## *Je wordt vanzelf een fan van Turing*

**Peter Grünwald** onderzoeker aan het Centrum voor Wiskunde en Informatica

“Toen ik in de twintig was, heb ik me veel beziggehouden met de vraag wat bewustzijn is. Dat was voordat ik de wiskundige kant op wilde. Bewustzijn willen begrijpen is de allermooiste en allerdiepste vraag, vond ik toen. Maar ik merkte ook dat we eigenlijk gewoon geen idee hebben wat het is.

“Dan word je dus vanzelf een fan van Turing. Als wij het verschil niet meer kunnen merken tussen mens en machine, dan kun je alsnog ruzie maken over of machines bewust zijn, want dat kun je nooit weten. Maar je kunt geen ruzie meer maken over de vraag of ze kunnen denken.

“Ik denk dan ook dat de Turingtest nog steeds relevant is, en ook dat er met grote taalmodellen een enorme stap gezet is om ‘AI-volgens-de-Turingtest-definitie’ te creëren. Daar zijn we zeker nog niet, en je kunt je ook afvragen of de Turingtest nou het beste criterium is. Maar deze modellen vertonen iets dat overeenkomsten heeft met intelligentie. Het is een soort denken, maar dan zonder karakter, gedrag, bewustzijn et cetera.

“Om mij heen hoor ik mensen vaak krampachtig verklaren dat ze ‘niet echt denken’, niet al-

tijd met de beste argumenten. Mensen willen zo graag dat deze modellen slecht werken. Maar ik gebruik ze zelf voor hulp in mijn wiskundig onderzoek en ik zie ze steeds indrukwekkender dingen doen. Ze maken stomme fouten, maar doen soms ook briljante suggesties. Net als collega’s aan wie ik iets vraag, maar dan één miljoen keer zo snel.

“Ik zou niet meer verbaasd zijn als een machine slaagt voor de Turingtest. Tien jaar geleden had ik gezegd dat wij dat niet meer mee gaan maken, maar ik vind dat er nu echt een megasprong is gemaakt. Ik snap ook eigenlijk niet goed hoe je met de gebruikte technieken zulke goede resultaten kunt krijgen. Dus als het nu zou gebeuren zou ik geen existentiële crisis krijgen.”

Met een grote lach: “Die heb ik nu al. Niet dat ik er heel ongelukkig van word, hoor. Die mensen die het over een existentiële dreiging voor de mensheid hebben, dat zijn vooral die AI-ontwikkelaars zelf, die dat leuk vinden om te zeggen. Volgens mij is dat een beetje over de top. Maar het is wel zo dat we iets gecreëerd hebben waarvan we gewoon totaal niet kunnen overzien wat het zal gaan doen. En dat maakt de wereld zo verschrikkelijk anders.”



Onder: wetenschapper Alan Turing in 1928, op 16-jarige leeftijd.



Boven: in de film 'The Imitation Game' speelt Benedict Cumberbatch Turing.

FOTO ANP