

Podcast – Life 3.0. | Max Tegmark

Vandaag gaan we in onze vertrouwde studio in Leuven een eerder filosofisch gesprek aan. Dat doen we met VLAIO bedrijfsadviseur Bruno Vande Castele naar aanleiding van het boek Life 3.0 van Max Tegmark over Artificiële intelligentie.

Werner Van Horebeek: Welkom Bruno

Bruno Van de Castele: Dag Werner

Werner Van Horebeek: Bruno, ik herinner me dat we een 6tal maanden geleden nog discussieerden of Artificiële Intelligentie een nieuwe winter tegemoet zou gaan. Sinds de term ergens in de 60'er jaren gemunt werd, zijn er wel een paar van die zogenaamde winters geweest. Periodes waarin het op de achtergrond sluimerde bij een aantal dia hards, weg van de public eye. Maar sinds het einde van vorig jaar en met de introductie van ChatGPT van OpenAI, is AI er helemaal terug van nooit weggeweest, zo lijkt het wel. Je kan geen krant of tijdschrift openslaan, geen podcast app of youtube openen of er wordt druk gedebatteerd, druk gespeculeerd over de mogelijkheden van AI nu het plots zo concreet en toegankelijk geworden is. Lange intro voor een eerder korte vraag: jij, die gepokt en gemazeld bent in de wereld van machine learning en NLP, hoe ervaar jij die plotse, recente aandacht voor AI? Is dat iets goeds, iets slechts?

Bruno Van de Castele: Dat is zeker iets goeds, en het toont ook het belang aan van het Vlaams beleidsprogramma rond Artificiële Intelligentie. De mogelijkheden van AI zijn groot, je kan ze ook niet meer wegdenken en niet meer zomaar afschaffen. Langs de andere kant, door z'n grote rekenkracht en eindeloze mogelijkheden kan AI ook leiden tot heel wat misbruik. We hebben samen een boek gelezen dat daar duidelijk ook voor waarschuwt.

Werner Van Horebeek: Aanleiding voor ons gesprek is het boek Life 3.0 van Max Tegmark. Daarin blikt hij vooruit in de behoorlijk verre toekomst om te verkennen wat AI voor onze wereld zou kunnen betekenen. Wat AI vandaag al vermag, wordt wel in kaart gebracht, maar de kaart van de mogelijkheden blijkt veel ruimer. Mag ik het een gedurfd boek noemen?

Bruno Van de Castele: Inderdaad, de auteur durft door te denken, vele jaren vooruit en naar wat er mis kan lopen met artificieel intelligentie. Vandaaruit koppelt hij terug naar het heden en wat je nu al kan doen om dat tegen te gaan. In dat opzicht is het ook een goeie tegenpool voor al wie het nu te rooskleurig inziet.

Werner Van Horebeek: Met de titel van het boek, Life 3.0, verwijst de auteur naar drie stadia in onze ontwikkeling: life 1.0, life 2.0 en life 3.0. Wat vind jij persoonlijk van die indeling, Bruno?

Bruno Van de Castele: Ja even toelichten, life 1.0 is waar hardware en software geëvolueerd zijn over vele jaren, miljoenen jaren. Dat is onze biologische achtergrond. Dat klinkt logisch. Life 2.0 is waarbij dat onderliggende biologische (hardware) geëvolueerd is, maar de software die erboven draait (bv onze cultuur, memes, ...) snel ontwikkelt en gedesigned kan worden. Evolutie duurt vele generaties, maar door die "software" aan te passen (onze cultuur, onze instellingen) zijn we veel flexibeler.

Dat klinkt al wat vreemder, maar life 3.0 (wat nu nog niet bestaat) is de volgende stap waarin zowel hardware als software ontwikkeld worden en dus veel sneller ontwikkeld kunnen worden. Susan Schneider, een andere denker over AI, formuleerde het zo: wij vliegen met een vliegtuig. Dat gaat op

een heel andere manier dan vogels met hun geëvolueerde vleugels, onze vliegtuigen zijn ontworpen en kunnen zo nodig, snel aangepast worden in een volgende versie.

Werner Van Horebeek: En wat is vervolgens de hoofdthese van het boek, Bruno?

Bruno Van de Castele: De hoofdthese van het boek "Life 3.0: Being Human in the Age of Artificial Intelligence" van Max Tegmark is dat de ontwikkeling van kunstmatige intelligentie ons leven op fundamentele manieren zal veranderen en dat we de verantwoordelijkheid hebben om ervoor te zorgen dat deze veranderingen ten goede komen voor de mensheid. Tegmark argumenteert dat we ons moeten voorbereiden op de komst van "Life 3.0", waarin AI de dominante vorm van intelligentie op aarde zal zijn en dat we er alles aan moeten doen om ervoor te zorgen dat AI ethisch verantwoord is en werkt voor het welzijn van de mensheid.

Werner Van Horebeek: Bruno, hold on for a second, moeten we hier geen disclaimer plaatsen?

Bruno Van de Castele: juist, ja, dat kunnen we maar beter doen. Het antwoord dat ik je zonet gaf over de hoofdthese van het boek komt niet van mij maar van Chat GTP, een generatieve tekst ai. We voerden een experimentje uit: we vroegen de performante chatbot eerst "welke vragen hij/zij/het zou stellen aan iemand die het boek "Life 3.0 gelezen had? We ontvingen een 10tal goede vragen terug. Die voedden we dan opnieuw aan ChatGTP en noteerden de antwoorden (plus selecteerden). Dus zowel de vraag als het antwoord hierboven kwamen niet van ons maar werden door ChatGTP gegenereerd op basis van de informatie die hij op internet vindt.

Werner Van Horebeek: En wat, als ik jou vragen mag, leert dat experiment jou over de stand van zaken in AI-wereld?

Bruno Van de Castele: Dat de lat hoger ligt, ook voor onze podcasts! We zijn het voorbije jaar echt getraakteerd op mooie evoluties qua generatieve AI (want zo noemt dat). Beelden, tekst, en binnenkort ook stem en video, AI kan het genereren, aan een grote snelheid en met verbluffende kwaliteit.

Maar, geen paniek Werner, creatieve jobs zoals de jouwe zijn nog niet weg. Wij als mensen hebben het idee gehad om die vragen te vragen, en om de antwoorden na te kijken en te selecteren. Daarvoor moesten we het boek toch lezen en moesten we onze ruime achtergrondkennis aanspreken. Het komt dus nog wel goed.

Werner Van Horebeek: Die hele ChatGPT heeft wel wat teweeg gebracht, anders. Het heeft alvast de awareness rond wat er mogelijk is aangewakkerd. AI gaan de reacties ook alle kanten uit. Sommige mensen zijn verbluft over wat er allemaal kan, anderen vinden het 'nog altijd maar niks', nog anderen uiten hun bezorgdheid in die zin, dat ALS dit nu al kan, WAT kan er dan allemaal binnen 5 of binnen 10 jaar? En daarmee zijn we bij de core van ons boek aanbeland: volgens de auteur houden we er maar best rekening mee dat AGI (Artificial General Intelligence) eraan staat te komen. Hoe sta jij in die discussie Bruno?

Bruno Van de Castele: Eerst misschien toelichten wat AGI, artificel general intelligence is. Het is een AI die "alles kan". Een tesla rijdt misschien nu wel al redelijk automatisch in een aantal omstandigheden, het kan niet zomaar overal rijden. Bovendien kan die tesla niet koffie zetten en ook niet de koffiekoeken zondag. Daarvoor kan je eventueel wel een andere AI trainen. De AI die we nu kunnen is vrij specifiek getraind op 1 taak (foto's herkennen, voorspellingen, ...), en is moeilijk over te zetten naar een compleet andere taak.

Versta me niet verkeerd. Onze huidige AI, de zgn “narrow ai” (smalle) kan al veel, maar moet ook nog flink getraind én opgevolgd worden. Vandaaruit de stap zetten naar AGI lijkt me te ver. Ik heb ook een analogie daarvoor: onze huidige AI is vergeleken met vroeger een ferrari ten opzichte van een 2 pk'tje. Het blijft evenwel een ding met 4 wielen, het is geen vliegtuig en ook geen vogel. Zelfs al blijven we AI verder ontwikkelen zoals het de voorbije jaren gaat, dan nog zie ik daar geen AGI uitkomen.

Werner Van Horebeek: Misschien moet je toch nog even kort toelichten wat het concept van AGI precies is en hoe het zich verhoudt tot wat we momenteel als AI kennen (foto herkenning, het maken van suggesties op basis van patronen, ...)? Is AGI dan geen probleem?

Bruno Van de Castele: Misschien wel. Ik kan niet 40-50 jaar vooruitkijken, misschien lukt het tegen dan na een paradigma-shift wel. Maar ik vind dat teveel focussen op AGI zoals Tegmark doet en zoals ook een aantal andere technologisten doen, de aandacht afleidt van het echte probleem dat we nu al hebben met AI misbruiken.

Tegmark zal mij klasseren als een techno-skepticus, maar zet ook diegenen die schrik hebben van de slechte zaken van een AGI weg als “luddites”. En dat is een beetje teveel op de man spelen, precies alsof een organisatie bv als AlgorithmWatch dat actief de huidige misbruiken nu aan de kaak stelt, achterlijken zijn. Ik heb vooral schrik dat door zo te focussen op een veel te verre en wellicht onrealistische toekomst, de huidige problemen niet aangekaart worden.

Werner Van Horebeek: Neemt natuurlijk niet weg dat we moeten (blijven) nadenken over de sociale, de IT-technische, de ethische en juridische uitdagingen waar AI ons voorplaatst, toch? Misschien eerst iets over de mogelijke sociale gevolgen? Recent publiceerden we met VLAIO een themadossier over Werkbaar Werk en bekeken hoe en waar technologie kan faciliteren om werk werkbaarder te maken. En ik hoor ook Satya Nadella, de CEO van Microsoft AI roemen om de mogelijkheid die het biedt om onze productiviteit te verhogen zonder nog harder of nog langer te moeten werken, alleen maar beter. Hoe sta jij in die discussie Bruno? Pakt AI ons werk af of maakt het ons productiever?

Bruno Van de Castele: AI heeft het potentieel om ons productiever te maken, daar zijn de Vlaamse bedrijven die wij als bedrijfsadviseur mogen begeleiden, het beste voorbeeld van. AI kan ons ook helpen om repetitieve of zelfs bullshit jobs (zie daarvoor het boek van David Graeber) te schrappen. Voor ons mentaal welzijn is dat zeker beter. En als AI ons kan helpen door wat nudging (cfr podcasts met collega Kris) te doen naar fysieke beweging, dan is dat ook goed.

Maar we moeten ons als maatschappij ook de vraag durven te stellen of we wel zoveel miljarden willen investeren in het verzamelen van data om ons dan zagezegd betere reclame advertenties te serveren, of om cv's te selecteren. Kunnen we die moeite dan niet beter steken in zaken die ons als maatschappij verder helpen? Een AI trainen heeft een grote kost aan energie, en dus ook CO2 uitstoot. We moeten daar spaarzaam mee zijn. Vandaag verscheen een bericht van NVIDIA (chip fabrikant) over hun “eye contact” software, waarbij de video van een persoon aangepast wordt zodat die altijd in de camera kijkt. Weet je wel hoe creepy dat is? Is dat echt nodig?

Kortom, technologie is goed, maar Tegmark merkt ook terecht op: we mogen er niet blind voor zijn.

Werner Van Horebeek: Interessante visie Bruno, technologie omarmen maar niet blindelings. Een andere uitdaging is het IT-technische aspect. Er zijn fysieke grenzen (zoals de snelheid van het licht, ...) die niet alles zomaar mogelijk maken. Jij had ook bedenkingen bij het Operating System van zo'n AGI en rond Cyber security en Cyberveiligheid is het laatste woord ook nog niet verteld. Ik smijt het

misschien allemaal op een hoopje, nu, maar er zijn nog wel wat IT-technische aspecten die bekeken moeten worden, toch?

Bruno Van de Castele: Natuurlijk, en ook vanuit de Vlaamse overheid worden die zaken aangemoedigd. Net zoals je computer moet je ook je AI beveiligen tegen piraterij of tegen misbruik. Daarvoor ook dat we hier in Vlaanderen een beleidsprogramma rond cybersecurity hebben. Voorbeelden te over van vroegere chatbots die opengesteld werden naar het publiek, en binnen de dag ranzige neonazi en white supremacy begonnen te spuien.

Tegmark gaat ook nogal snel voorbij aan twee belangrijke aspecten binnen IT. Ten eerste draait die op computerchips, en die moeten ontworpen worden. In Vlaanderen hebben we IMEC, een wereldbekende onderzoeksinstituut voor ontwerp van computerchips. Ook beveiliging wordt daarin actief meegenomen. En ten tweede, ja elk algoritme heeft een besturingssysteem nodig, en wie 5 minuten met een computer werkt, weet dat dat niet zo vanzelfsprekend is om te maken, te gebruiken en te beveiligen. Ook daar is dus nog flink wat werk.

Werner Van Horebeek: Ethiek is gelukkig ook nooit veraf in de rijke discussie over Artificiële Intelligentie. Kan je dat even onder de loep nemen?

Bruno Van de Castele: In deze podcastserie spraken collega's Kris Honraet en Celina Buysse daar ook al over. Misschien kan je de links naar die podcasts ook in de liner notes zetten?

Werner Van Horebeek: Goede suggestie, Bruno.

Bruno Van de Castele: Bij ethiek gaat het over biases die de data kunnen hebben en die de AI blind overneemt, het gaat om de AI die als een black box opereert en antwoorden genereert waarvan we niet meer kunnen zien waar ze vandaan komen. ChatGPT, die we hierboven al gebruikten als co-presentator, vertelt soms met heel veel vertrouwen de meest foutieve zaken. We moeten dus als sector nog meer werk steken in die betrouwbaarheid te verhogen.

En dat naast natuurlijk het opkuisen van data om ermee te kunnen trainen en om die bias daar expliciet uit te halen. Daar steek je als AI ontwikkelaar heel veel tijd in, het is niet het meest toffe aan de job, maar het moet wel gebeuren. Dat is ook gebeurd met ChatGPT, en is nu al bezig met de volgende versie – ja want dat duurt soms meerdere jaren om te doen.

Werner Van Horebeek: Waarmee we meteen bij het juridische zijn aanbeland.

Bruno Van de Castele: Klopt, we moeten ons de vraag stellen : willen we echt al onze data zomaar prijsgeven? GDPR en verder, bescherming van het individu en haar privacy. We zijn daar in Europa terecht zeer voorzichtig mee, en je ziet toch wel op wereldniveau dat het evolueert. Het is niet omdat het op internet staat, dat we dat zomaar kunnen gebruiken. Ik verwijs hiervoor graag naar het werk dat aan onze vlaamse universiteiten gebeurt, bv Ann-Katrien Oimann die ook actief mee bouwt aan die multi-disciplinaire reflectie over wat we met die data in Europe willen doen.

Werner Van Horebeek: Een boeiende reis vond ik de lectuur van dit boek wel. Positief vond ik alvast de wetenschappelijke approach met heldere definities van duistere begrippen als intelligentie en intelligentie explosie. Ook het denken in scenario's beviel me wel. Maar ik begreep dat jij ook wel wat wetenschappelijke kritiek bij het boek had?

Bruno Van de Castele: Goh ja, ik heb toch wat opmerkingen die een fundamenteel probleem aanduiden. Zo vind ik het vreemd dat Tegmark als fysisch er zomaar van uitgaat dat we onze hersenen kunnen "opladen" in een siliconen-substraat om dan verder eeuwig te leven. Volgens het

Heisenberg principe is dat quasi onmogelijk om te doen in één keer voor onze vele miljarden neuronen.

Daarnaast mist Tegmark toch wel een multi-disciplinair inzicht. Als hij wil kan ik hem Top-Vlaamse instellingen aanraden om mee in conversatie te gaan. Het is vreemd dat hij het heeft over de Experience machine, maar niet vermeld dat deze van Nozick is, een filosoof. Hij spreekt soms wel eens over “de filosofen” maar citeert er geen enkele met naam, wat hij wel voor andere zaken doet. Dat geeft aan dat hij niet helemaal mee is met waar andere wetenschappers mee bezig zijn.

Zo spreekt hij ook over een naïef onderscheid tussen hardware en software als het gaat over hersenen en bewustzijn. Er zijn wellicht luisteraars van de Nerdland podcast van Lieven Scheire onder de luisteraars (een aanrader trouwens), en wel in elke aflevering wordt er bijna geroepen dat je voor onze hersenen niet mag spreken in zo’n kunstmatige opsplitsing. Opnieuw dus een lacune over de huidige wetenschappelijke kennis, en dat zou in dit dikke boek toch niet mogen.

Werner Van Horebeek: Wat betekent die kritiek voor je algemene appreciatie van dit boek, Bruno. Het is al bij al een stevig boekwerk van bijna 500 bladzijden. Wie zou je het aanbevelen om te lezen?

Bruno Van de Castele: Wel, als wetenschapper of IT’er ga je snel afknappen op die lacunes. Maar versta mijn kritiek niet verkeerd: als ondernemer of voor het brede publiek is dit een goeie kennismaking met wat AI is, zowel narrow als general. Zeker de eerste helft van het boek is erg didactisch geschreven, en dus de moeite waard.

Bovendien moet je ook de tweede helft lezen. De bewustwording dat we niet zomaar blind op technologie moeten vertrouwen is heel belangrijk, al kon het met wat minder bladzijden. Maar als je maar 1 ding onthoudt, bv als ondernemer die met AI wil werken: ja, zeker, doen, er zijn vele positieve mogelijkheden aan. Maar wees je bewust van die beperkingen, stap niet mee in misbruik en onnutte zaken. Je kan er nu al mee aan de slag.

Werner Van Horebeek: heel erg bedankt voor jouw visie, Bruno. Als we in de toekomst nog moeilijke onderwerpen te tacklen hebben, weten we bij wij we moeten zijn. AI dan niet geholpen door ChatGTP.