

Hier ligt een ethische taak voor de manager

HET AMORELE VAN AI MET (SOMS)

IMMORELE GEVOLGEN

ARTIFICIËLE INTELLIGENTIE (AI) ZOALS DIE NU IN ORGANISATIES WORDT TOEGEPAST, IS ZEER EFFECTIEF EN EFFICIËNT IN CONTEXTEN DIE RELATIEF 'STABIEL' ZIJN. TOT ENIGE ETHISCHE REFLECTIE IS HET ECHTER NIET IN STAAT. MANAGERS MOETEN ZICH DAAROM VOLGENS PETER VERHEZEN REALISEREN DAT ZIJ ZELF VERANTWOORDELIJK BLIJVEN VOOR HET ETHISCH GEDRAG VAN HUN AI-SYSTEMEN.

door Peter Verhezen beeld Shutterstock

DE ALGORITMEN ACHTER DE SUGGESTIES DIE NETFLIX, GOOGLE, WEIBU, FACEBOOK OF AMAZON EN ALIBABA ONS AL DAN NIET GEVRAAGD AANBIEDEN, WETEN ERG ACCURAAT TE VOORSPELLEN WAT MENSEN VERLANGEN, OF DAT NU FILMS OF BOEKEN BETREFT.

Maar wanneer we met grotere onzekerheden en turbulentie te maken krijgen, zal de voorspellingskracht van AI van minder waarde zijn, omdat de historische data achterhaald zullen zijn en/of 'onbetrouwbaar'. De geschiedenis garandeert niet dat het verleden gelijk zal zijn aan de toekomst. En kritische normatieve vragen worden amper gesteld in de programmeerwereld van algoritmen, en al evenmin hoe data consciëntieus kunnen worden gebruikt.

Paradoxaal genoeg is artificiële intelli-

gentie eigenlijk niets anders dan "(dumb) computers on steroids".

DE GROTE BEPERKTHEID VAN AI

AI mag een schaakspel kunnen winnen van een grootmeester, ons brein blijft voorlopig superieur aan de beste AI. De menselijke superioriteit heeft te maken met drie factoren die AI voorlopig niet binnen handbereik heeft: computers kunnen hun informatie niet op een flexibele manier communiceren; evenmin kunnen ze zich automatisch aanpassen aan een totaal verschillende of veranderde context; en ten derde heeft de mens een autonomie om in te grijpen en zich over de toekomst te verwonderen, wat AI (voorlopig nog) niet kan. Dit laatste heeft ook met ons ethisch vermogen te maken, wat een algoritme totaal niet heeft. Een 'intelligente'

AUTEUR



PETER VERHEZEN is hoogleraar (Global Strategy & Emerging Markets) op de Antwerp Management School (AMS), en principal van Verhezen & Associates (Corporate Governance & Risk Management advisory in Asia), www.verhezen.net.

ONS BREIN
BLIJFT VOORLOPIG
SUPERIEUR AAN
DE BESTE AI



computer kan dus (nog) niet reflecteren over zichzelf.

Kunstmatige intelligentie begrijpt niets van wat een mens zoal kan doen of denken. Denken in causale modellen is een unieke, menselijke vaardigheid die ons in staat stelt om gemakkelijk bepaalde kennis naar andere contexten over te dragen. Intelligentia in silicone is wel in staat om zelf te 'leren' door middel van feedbackmechanismen, en op die manier schijnbaar autonoom te 'denken'.

Maar een slim algoritme begrijpt helemaal niet waarom die aanpassing nodig is. En naarmate deze self-learning algoritmen steeds complexer en efficiënter worden – onderbouwd door convolutional network layers die meer en meer tot een 'commodity' verworden – verworden ze al snel tot een 'black box', waarbij gesuggereerde beslissingen totaal niet transparant zijn, of zelfs niet meer helemaal begrepen worden door de programmeurs zelf.

De manier waarop algoritmen achter machine learning worden geprogrammeerd – zelfs in het geval van 'autonoom leren van AI' – kan veelal tot

Bayesiaanse 'what if, then'-logica worden teruggebracht, wat betekent dat ze zeer goed probabilistisch 'denken'. Deze vorm van artificiële intelligentie en zijn data heeft echter geen notie van ethiek of waarden die door programmeurs of managers stilzwijgend of onbewust worden vóórondersteld. En dus geldt ook hier: garbage in, garbage out. Zonder een kritisch ethische reflectie – die limieten stelt aan deze algoritmen – kan dit snel resulteren in ethisch ongewenste resultaten. Microsofts self-learning bot was zo'n voorbeeld dat de 'kennis' snel en efficiënt kon verzamelen van social media, maar vergeleed tot een racistisch

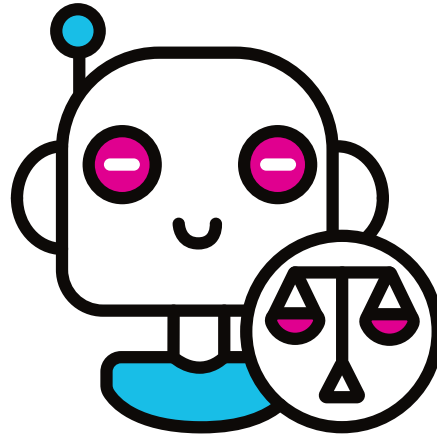
en rauw medium dat onaanvaardbaar werd geacht omdat deze chatbot geen ethisch onderscheid kon maken van wat contextueel al dan niet aanvaardbaar was. Microsoft had geen andere keuze dan deze technologie niet verder commercieel aan te bieden omwille van het onethische karakter van het AI-product.

MECHANISTISCHE 'ZIELLOZE' NATUUR

Kunstmatige intelligentie heeft geen bewustzijn en is evenmin in staat zelfstandig na te denken over ethische vraagstukken. Patronen die in big data worden gevonden, mogen dan tot nieuwe inzichten leiden, we mogen niet verwachten dat algoritmen objectief neutraal zouden zijn. Black boxes zijn niet in staat (ethische) intenties te creëren. Door hun mechanistische 'ziellose' natuur zijn algoritmen neutraal of amoreel. Maar dat betekent niet dat ze geen negatieve ethische gevolgen kunnen hebben: algoritmen met betrekking tot preventive policing in Chicago, of automated hiring bij Amazon hebben aangetoond dat dit tot erg onethische gevolgen kan leiden.

AI blijft een niet bezielde mechanisme zonder enige vorm van begrip. De enige vorm van objectiviteit is een statistisch naar het 'midden' neigende 'groepswijsheid'. Big data stellen ons in staat om een zekere nivellering naar het gemiddelde tot stand te brengen, iets wat een individu per definitie moeilijker kan realiseren omwille van de vele onbewuste subjectieve vooroordelen (of biases). Voorspellingsexperimenten hebben aangetoond dat zelfs experts niet veel beter presteren dan een chimpansee of het werpen van de dobbelsteen (randomness).

Snelle computers en artificiële intelligentie weten op een bewonderenswaardige efficiënte en minutieuze manier correlaties te leggen – het vinden van patronen in big data – die wij mensen nauwelijks kunnen waarnemen. Omdat computers puur statistische calculaties maken, en helemaal niet kunnen 'den-



ken', kunnen ze ook geen 'bewuste' causale verbanden leggen. AI vindt statistische correlaties die zich visualiseren in datapatronen die zelfs experts niet altijd zien. AI kan echter oorzaak en gevolg niet op een objectieve wetenschappelijke manier onderscheiden. En zonder causale verbanden te kunnen leggen – iets wat wij continu doen – is ook ethische reflectie onmogelijk.

INZETTEN OP HUMAN-ASSISTANCE SYSTEMS

Dat bepaalde herhaalbare monotone taken worden geautomatiseerd en sneller kunnen worden uitgevoerd door smart computers of robots, is nu reeds een vanzelfsprekendheid. Een mens daarentegen is in staat om creatief en ondernemend handelingen te treffen waarvan een computer alleen maar kan dromen – iets wat een algoritme evenmin kan natuurlijk. Techno-chauvisme – het idee dat wiskundige formules ingebed in een code beter in staat zijn maatschappelijke of businessproblemen op te lossen dan mensen – biedt niet echt oplossingen voor onze sociaal-maatschappelijke en ethisch-maatschappelijke problemen. Juist omdat mensen kritisch en ethisch nadenken over technologie en over wat we voornog al dan niet kunnen automatiseren, mogen we dit niet aan programmeurs overlaten, maar adviseren we een nauwe samenwerking tussen technici en

kritisch denkende managers en ethici die samen bepalen wat kan en niet kan, wat geautomatiseerd kan worden en wat niet. Algoritmen en hun black boxes moeten steeds in vraag worden gesteld, en ethische limieten moeten duidelijk worden gesteld, zodat we niet in ethisch onaanvaardbare situaties terechtkomen die we niet hadden voorzien. We moeten inzetten op 'human-assistance systems' in plaats van 'human-replacement systems'. Zoals alle wetenschap, zal ook technologische innovatie van AI-producten zich via trial-and-error verder ontwikkelen.

En al is AI-technologie moreel neutraal, het Facebook-debacle in 2016 en recenter heeft bewezen dat heel wat data op een weinig ethische manier kunnen worden misbruikt. Het recente ontslag van twee toponderzoekers die kritiek hadden op de onethische aspecten van Googles strategie rond kunstmatige intelligentie maakt duidelijk dat ethiek regelmatig onder de mat wordt geveegd. Naast 'guidelines for trustworthy and responsible AI' en de recente EU-GDPR-regulering – 'protection of privacy of data' – en andere normen, zijn het de bestuursraden en de executives die verantwoordelijk worden geacht voor het ethisch en legale gedrag van hun smart computers. Maar zelfs de GDPR loopt nu reeds achter over hoe we de privacy van data in de AI-black boxes kunnen garanderen.

Bestuursleden en executives blijven uiteindelijk verantwoordelijk voor de reputatie van hun onderneming. Hun beslissingen hebben veelal ethische gevolgen; vooral omdat heersende waardebeoordelingen en normen impliciet of expliciet worden 'geobjectiveerd' in algoritmen die gedreven worden door de 'nieuwe olie' in business, namelijk big data. Het is de taak van het hoger onderwijs om (toekomstige) business-leaders bewust te maken van zowel de negatieve als positieve impact van AI-technologie. Een kritisch ethische bevraging met betrekking tot AI zal ons alleen maar ten goede komen. 🌐