



'Technologie is slechts een hefboom op onze eigen intenties. Laten wij daarom in godsnaam nog even aan het stuur blijven.'

FOTO: HH

Het debat over de ethiek van kunstmatige intelligentie wordt vooral gevoerd in het licht van data verzamelen en data verwerken. Zo waarschuwt belangenorganisatie Bits of Freedom al langer voor zogenoemde datasleepnetten. Dat zijn doelbewuste acties van overheden en bedrijfsleven om zoveel mogelijk informatie te verzamelen, in plaats van alleen wat relevant is voor de situatie.

Verfrissend is dan bijvoorbeeld de benadering van De Correspondent. Zij hanteren een 'dataluw' beleid: om de verhalen van hun journalisten te lezen hoeft je slechts je naam en e-mailadres op te geven. Aanvullende gegevens zoals geslacht, geboortedatum en postcode zijn niet interessant omdat De Correspondent geen adverteerders heeft om deze data mee te delen.

Alle organisaties in de publieke en private sector zouden hierover beleid moeten formuleren, al was het alleen al omdat in mei 2018 de Wet bescherming persoonsgegevens vervangen wordt door de Algemene verordening gegevensbescherming (AVG). In de hele EU geldt dan een nieuwe privacywetgeving.

Naast het verzamelen van data zou ook het verwerken van data niet vrij moeten zijn van ethische reflectie. Sinds 2016 werkt een collectief van Nederlandse universiteiten aan 'Responsible Data Science'. De onderzoekers adviseren bij het verwerken van data de volgende vier criteria van het Fact-principe in acht te nemen:

**1 Fairness (redelijkheid).** Voorkom dat groepen die slecht vertegenwoordigd waren bij het verzamelen van de data nadeel ondervinden van de uiteindelijke werking. Zo zal Google er nog jaren aan worden herinnerd dat bij de eerste versies van hun fotoherkenningssoftware de selfies van mensen met een donkere huidskleur per ongeluk werden gecategoriseerd als gorilla's.

**2 Accuracy (zorgvuldigheid).** Als je maar genoeg hypothesen onderzoekt, is er altijd wel eentje die per ongeluk waar lijkt te zijn. Een hilarisch voorbeeld is te vinden in de Amerikaanse staat Maine, waar het aantal echtscheidingen bijna perfect correleert met de verkoop van margarine. Dát er een verband lijkt te zijn wil nog niet zeggen dat je er een conclusie aan mag verbinden.

**3 Confidentiality (vertrouwelijkheid).** Hoe kun je garanderen dat privacygevoelige informatie wordt gerespecteerd? Als in 2018 de AVG van toepassing wordt, kunnen sommige bedrijven verplicht worden een privacy-impactassessment uit te voeren of zelfs een functionaris voor de gegevensbescherming aan te stellen.

**4 Transparency (openheid).** Is de manier waarop je tot een uitkomst komt verdedigbaar en inzichtelijk? Als mensen zich benadeeld voelen door jouw algoritme, kunnen ze inzicht eisen in hoe het algoritme werkt. (Zie kader 'Een algoritme als rechter'.)

Deze vier criteria vormen een goed fundament voor een nieuw privacy-state-

**Dilemma**  
**Wij kunnen de verantwoordelijkheid voor onze daden niet afschuiven op ons gereedschap**

## Algoritme straft mens

In de Verenigde Staten stond Eric Loomis terecht voor een schietpartij. Hij werd schuldig bevonden en vulde op verzoek van de rechter een lange digitale vragenlijst in. De rechter gaf Loomis vervolgens een relatief lange straf op basis van de uitkomst van de software.

Toen Loomis hiertegen in verweering, eiste hij inzage in de motivatie voor zijn hoge straf. In de zaak Wisconsin versus Loomis werd hij door het hooggerechtshof van de staat in het ongelijk gesteld. Loomis zal de zaak waarschijnlijk doorzetten naar het federale Hooggerechtshof. Deskundigen wijzen erop dat te veel vertrouwen op software risico's met zich meebrengt op het gebied van veiligheid en justitie. Zolang zulke algoritmen opereren als een 'black box' (niet transparant), kan een rechter geen verantwoordelijkheid nemen voor het besluit van de software.

ment, of het inrichten van een 'privacy board'. Iets wat bij ING al is gebeurd. De bank heeft geleerd van de commotie die enkele jaren geleden ontstond bij een proefballonnetje rond big data, en heeft een stuurgroep in het leven geroepen die één keer per maand bij elkaar komt om nieuwe projectvoorstellen te beoordelen op ethische aspecten en die ook het mandaat heeft om voorstellen af te wijzen.

Maar met het in acht nemen van de privacyregels en het opzetten van een verantwoord dataverwerkingsbeleid hebben we het probleem van de morele dilemma's nog niet opgelost. Tech-optimisten geloven dat 'het goede doen' in de toekomst meegeprogrammeerd kan worden in de software. Ze noemen dat 'ethical by design'. Hoezeer ik mezelf ook onder de tech-optimisten schaar, denk ik dat die toekomst nog erg ver weg is. Van de honderden onderzoekers naar kunstmatige intelligentie in Nederland is er slechts een handvol bezig met de grote vragen rondom ethiek.

Data-science wordt wel eens vergeleken met een raket. De data zijn in die vergelijking de brandstof en de kunstmatige intelligentie is de motor. Alle bijbehorende processen kun je op een veilige manier — al dan niet met de vier Fact-principes — inrichten. Maar waar de raket naartoe moet vliegen, dát is de echte vraag. De richting bepalen vereist een morele afweging.

De grote vragen van dit moment verdienen een antwoord: een juiste richting. Net als dat bedrijven ethische commissies instellen, pleit ik voor een nationale raad van wetenschappers. En niet alleen vertegenwoordigers van de afdelingen informatica en data science, maar juist ook filosofen, juristen en ethici.

We kunnen deze discussies niet voor ons uit blijven schuiven. Hoe meer we de tekortkoming van machines leren kennen, hoe beter we onszelf leren kennen. Het is een misverstand dat technologie neutraal zou zijn. Dat is niet zo. Technologie is slechts een hefboom op onze eigen intenties. Laten wij daarom in godsnaam nog even aan het stuur blijven zitten en eerst de best mogelijke richting van onze toekomst bepalen.

**Jim Stolze** is tech-ondernemer op Amsterdam Sciencepark. Daarvoor was hij de drijvende kracht achter de TEDx-conferenties in Europa en het Midden-Oosten.

# Techniek is niet neutraal: houd het stuur in handen

Jim Stolze

Omdat data een steeds grotere rol spelen in ons leven, wordt het hoog tijd om werk te maken van ethische reflectie. Jim Stolze pleit daarom voor een nationale raad van wetenschappers om de juiste richting te geven aan digitalisering.

**H**et is woensdagmiddag. Kort na de lunch gaan Jip en Annelies (een tweeling van 8 jaar) de deur uit om de hond uit te laten.

Een stuk verderop stuurt een zelfrijdende auto met een ouder echtpaar op de achterbank diezelfde straat in. De man van 89 heeft kort geleden te horen gekregen dat hij ongeneeslijk ziek is en heeft een afspraak met zijn behandelend arts.

Jip en Annelies zien het voetgangerslicht op groen springen en huppelen het zebrapad op. Op hetzelfde moment krijgt de zelfrijdende auto een melding dat het remsysteem is uitgevallen. In een fractie van een seconde overweegt het besturingssysteem twee opties: (1) Door rood rijden met de dood van de twee kinderen tot gevolg, of (2) Het stuur omgooien en de auto in volle vaart tegen een muur laten botsen, wat de twee inzittenden hun leven zal kosten.

Wat vindt u dat de auto zou moeten doen? De meeste deelnemers aan dit gedachte-experiment vinden dat het leven van de kinderen voorrang zou moeten krijgen.

Minder uitgesproken zijn de onder-vraagden wanneer we kleine onderdelen binnen het experiment aanpassen. Zou

u dezelfde mening zijn toegedaan wanneer de kinderen de verkeersregels hadden overtreden door rood te lopen? Of wanneer het slechts één kind betrof, en de auto vol zat met vijf kwieke senioren op weg naar hun wekelijkse yogaklasje?

Met dit soort morele dilemma's houden onderzoekers van het MIT Medialab zich bezig. Zij proberen inzicht te krijgen in menselijke besluitvorming om zo beter in te kunnen spelen op de verregaande digitalisering van de maatschappij.

Immers, we besteden steeds meer onderdelen van ons leven uit aan machines. Software is stilletjes uit onze computers geslopen en klimt via onze vingers en armen naar onze hersenen. Daarmee neemt software niet alleen over wat we doen, wat we denken, maar ook wát we hiervan zouden moeten denken.

Het gevaar dat op de loer ligt, is dat we zo ook de verantwoordelijkheid voor onze daden uitbesteden. Dat voelt als een brug te ver. Prima dat slimme machines de arbeidsmarkt komen versterken, en gezellig dat we robots ook op straat tegen kunnen komen, maar wij — de mensen — kunnen de verantwoordelijkheid voor onze geautomatiseerde daden niet afschuiven op ons gereedschap, ook al vertoont dat gereedschap menselijke trekjes.