

We hebben deugden nodig

Het gaat steeds vaker over algoritmes en artificiële intelligentie (AI) in overheid en openbaar bestuur. Zo was daar het 'Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen' van de Raad van State, over geautomatiseerde besluitvorming, zelflerende systemen en ketenbesluitvorming. Een mooi onderwerp. Controversieel ook. In de context van bestuur wil je immers ook in kunnen grijpen in automatische besluitvorming en beslissingen kunnen traceren.

Er spelen allerlei technische, economische, juridische en ethische issues mee. De technische en economische issues kunnen we samenvatten als: er kan steeds meer, en het kan steeds goedkoper en makkelijker. Marties van Eck onderzoekt de juridische issues van 'geautomatiseerde ketenbesluiten' ('Burger onvoldoende juridisch beschermd'). En Marriette Schaake pleitte voor betere wetgeving ('Kunstmatige intelligentie hoort in de wet'). Allemaal zeer nuttig en hard nodig. Wat betreft de ethische issues echter, is er slecht nieuws: er zijn geen 'kant-en-klare methoden' om ethische dilemma's op te lossen ('Hoe houd je de balans tussen AI, big data en ethiek?'). Maar gelukkig is er ook goed nieuws: professionals kunnen leren om ethisch om te gaan met algoritmes en AI door het cultiveren van deugden zoals zelfbeheersing, bescheidenheid, rechtvaardigheid, moed en perspectief.

DEUGDETHIEK

Ik pleit voor de inzet van deugdeethiek. Deze ethische traditie is gericht op het creëren van een rechtvaardige samenleving waarin mensen kunnen floreren, en die mensen aanzet om deugden te cultiveren die daaraan bijdragen. Het cultiveren verwijst naar het ontwikkelen van gevoelens, gedachten en handelingen in de richting van rechtvaardigheid en van floreren. Het gaat om het op één lijn brengen van voelen, denken en handelen. Je voelt dat iets onrechtvaardig is, je zoekt uit hoe je dit kunt rechtzetten, en je onderneemt actie.

Deugdeethiek is ook een manier om de menselijke maat



voor ethische algoritmes

terug te brengen ('De computer kan het ook fout hebben'). Het gaat namelijk niet over regeltjes, maar over improviseren; steeds vanuit een specifieke context. Neem 'moed', een klassieke deugd. Als ik een oude man ben, niet zo goed ter been, en ik zie een vechtpartij op straat, dan is moed voor mij: mezelf in veiligheid brengen en 112 bellen; ik zou overmoedig handelen als ik tussenbeide zou komen. Maar als ik een sterke kerel ben en getraind in vechtsport, dan betekent moed voor mij: rustig ingrijpen en een omstander vragen om 112 te bellen; aan de kant blijven staan zou laf zijn. Het gaat steeds om het vinden van het juiste midden.

In haar boek *Technology and the virtues: A philosophical guide to a future worth wanting*, stelt Shannon Vallor, hoogleraar aan Santa Clara University in Silicon Valley, voor om deugdeethiek te gebruiken bij het ontwikkelen en toepassen van nieuwe technologie. Ze legt uit dat de andere bekende ethische tradities (gevolgenethiek en plichtethiek) niet werken voor nieuwe technologie. Voor gevolgenethiek moet je precies alle voor- en nadelen kennen en tegen elkaar afwegen. Voor plichtethiek moet je zoeken naar universele regels die in alle omstandigheden zouden moeten gelden. Maar dat is niet mogelijk bij nieuwe technologieën, precies omdat die nog niet uitgekristalliseerd zijn; je kent de voor- en nadelen niet in detail en je kunt niet anticiperen op alle omstandigheden.

Laten we enkele relevante 'technomorele' deugden kort onder de loep nemen (uit *Technology and the Virtues*).

1. Zelfbeheersing

Mensen die aan algoritmes werken kunnen zelfbeheersing cultiveren. De verleiding kan groot zijn om allerlei data van allerlei instanties te verzamelen, aan elkaar te koppelen, en daar allerlei analyses op los te laten. Zo'n gulzigheid kan echter uitdraaien op een puinhoop en op illegale praktijken. Zelfbeheersing kan helpen om te streven naar minimaal gebruik van data en naar begrijpelijke algoritmes. Bijvoorbeeld een algoritme dat alleen data vanuit de eigen organisatie gebruikt (niet van andere instanties) en een algoritme zoals 'als groter dan X, dan A' (geen algoritmes die je niet kunt begrijpen en dus niet kunt uitleggen).

2. Bescheidenheid

Een andere deugd die kan helpen is bescheidenheid. Dat gaat over het weten wat we niet weten; over de tekortkomingen van algoritmes. Je weet bijvoorbeeld dat er altijd false positives en false negatives uit een algoritme kunnen komen, en je houdt daar rekening mee. Je kunt bescheidenheid cultiveren door een proces zo te ontwerpen dat mensen in control zijn voor het interpreteren van twijfelgevallen, of door het algoritme uitleg te laten geven bij de uitkomsten, bijvoorbeeld door het visualiseren van onzekerheid. Je neemt de zogenaamde waarheid van een algoritme met een korreltje zout.

3. Rechtvaardigheid

Rechtvaardigheid gaat over het streven naar een eerlijke

Door **Marc Steen**
Beeld **Shutterstock**

Luisteren, innoveren en leveren

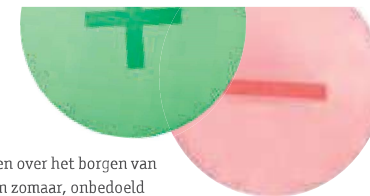
Samen met energieleverancier Eneco verduurzamen we de Rotterdamse warmtevoorziening door middel van blockchain. Ook dat is digitaliseren – onze kracht.

Luisteren, innoveren, leveren. Zo voorzien we onze opdrachtgevers effectief van een moderne, datagedreven en veilige bedrijfsvoering – en daarmee van een duurzame digitale toekomst. Luisteren om klant- en marktinzichten op te doen. Resultaatgericht innoveren door in co-creatie slimme oplossingen en onderscheidende dienstenconcepten te ontwikkelen. En leveren zoals van ons verwacht mag worden: end-to-end, de best passende adviezen en producten, met maximale waarde voor opdrachtgevers, hun klanten en burgers. CGI, al ruim 40 jaar digitaal leiderschap.

cginederland.nl

CGI

Experience the commitment®



Podium

verdeling van voordelen en nadelen, en over het borgen van rechten, zoals waardigheid. Er kunnen zomaar, onbedoeld onrechtvaardigheden binnensluipen in een algoritme. Een berucht voorbeeld speelde in 2015: Google lanceerde toen image recognition gebaseerd op machine learning, en labelde twee donkere tieners als 'gorilla's'. Dat kwam doordat er alleen foto's van lichtgekleurde mensen waren gebruikt in de trainingsdata. Dat hadden ze kunnen voorkomen door de trainingsdata een betere afspiegeling te maken van de bevolking.

4. Moed

Dit betoog gaat over het sturen van technologie, niet het tegenhouden van technologie. De deugd van moed gaat over het balanceren tussen hoop en angst, over het vinden van het juiste midden tussen blind vertrouwen in technologie en ongegronde hoop aan de ene kant, en klakkeloos afwijzen van technologie en ongegronde angst aan de andere kant. Concreet kunnen mensen de deugd van moed cultiveren door dingen zorgvuldig uit te proberen, bijvoorbeeld in gecontroleerde experimenten, en kritisch te volgen wat de gevolgen van zo'n experiment zijn, zowel de positieve als de negatieve, en om dat continu bij te sturen.

5. Perspectief

Ook de deugd van perspectief is nodig, en dan met name het combineren van verschillende perspectieven. Het ene moment zoom je in op het algoritme, bijvoorbeeld voor predictive policing, en probeer je biases in het algoritme op te lossen. Het andere moment zoom je uit naar een niveau van de maatschappij, en kijk je naar verschillende kansen voor verschillende groepen in de samenleving en de effecten daarvan op criminaliteit. Ook kun je perspectief wisselen vanuit verschillende vakgebieden: hoe kijken mensen naar dit algoritme vanuit data science, vanuit bestuursrecht, en vanuit kosten-baten.

PROFESSIELE ETHIEK

Deugdetiek is bij uitstek een professionele ethiek: een ethiek die inspiratie en richting en handvatten kan bieden aan het voelen, denken en handelen van professionals.

Soms vragen mensen mij: hoe operationaliseer je deugdetiek? Je kunt deugdetiek niet vangen in getallen die je kunt optellen, of in regels die je kunt opschrijven en die altijd gelden. Je kunt echter deugdetiek prima operationaliseren door er concreet mee aan de slag te gaan.

PRAKTISCH AAN DE SLAG

Stel dat je bezig bent met het ontwerpen of implementeren of gebruiken van algoritmes en AI in overheid en openbaar

Zelfbeheersing kan helpen om te streven naar minimaal gebruik van data en naar begrijpelijke algoritmes

bestuur. Dan kun je starten met een kort zelfonderzoek: welke van deze deugden heb ik nodig? Zelfbeheersing? Bescheidenheid? Rechtvaardigheid? Moed? Perspectief? Welke van die deugden wil ik cultiveren om mijn werk beter te kunnen doen?

Daarna kun je manieren vinden om één tandje beter te worden in één van die deugden. Je kunt eens 'nee' zeggen tegen een nieuwe feature. Je kunt de tekortkomingen van algoritmes meer aandacht geven. Je kunt aan de bel trekken als je vermoedt dat de toepassing van een algoritme rechtvaardigheid onderuit haalt. Je kunt een experiment bedenken en opzetten. Je kunt schakelen tussen inzoomen op het algoritme en uitzoomen naar de maatschappij.

En als dat goede resultaten heeft, dan doe je het nog een keer. Zo cultiveer je deugden. Een andere manier om deugden te cultiveren is van anderen leren. Daarom is het ook altijd nodig om successen te vieren en leerervaringen te delen.

Marc Steen is senior onderzoeker bij TNO