

Computer doet niet aan gelijke behandeling van mensen

Raad van Europa waarschuwt: regels zijn nodig om discriminatie aan te pakken

Jan Fred van Wijnen
Amsterdam

Stel je voor: je solliciteert op een nieuwe baan. De werkgever vraagt of je niet liever dat hele gedoe overslaat: al die vragen over je achtergrond, hoe lekker je gemotiveerd bent om het verschil te maken, wat je ideeën zijn over vleeseters. Je geeft gewoon even toestemming om je e-mails door te spitten, en de computer maakt er een 'persoonlijkheidsprofiel' van.

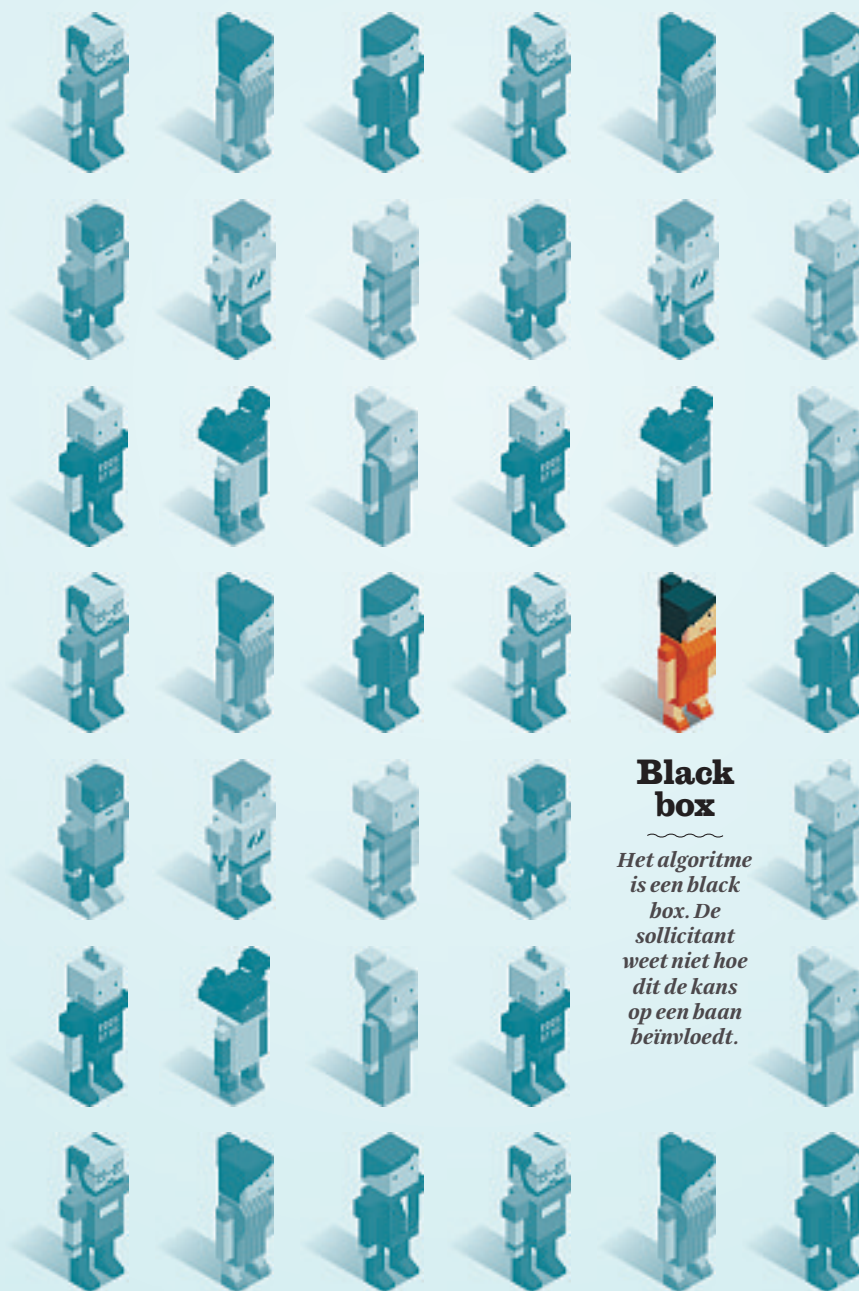
Zo gek ben je toch niet? Kennelijk wel. De kleine Finse consultant Digital Minds biedt deze service aan bij bedrijven. Volgens een recent rapport van Algorithm Watch, een Europese onderzoeksgroep, heeft nog geen kandidaat geweigerd.

Strikt genomen wordt de Europese privacywet GDPR gerespecteerd: je moet eerst braaf de inloggegevens voor je e-mailaccount overhandigen. Het is wel de vraag hoe vrijwillig dat is, voor iemand die heel graag een baan wil.

Reijo Aarnio, de Finse ombudsman voor databescherming, onderzoekt de zaak. Hij geeft nog geen commentaar, schrijft hij per e-mail, want 'de inspecteur die op de zaak zit, denkt nu na over zijn volgende stap'.

Het Finse bedrijf zelf reageert niet op een informatieverzoek van

**'Slimme' systemen
rukken op bij politie,
bedrijven en banken,
en de wetgeving laat
te wensen over**



Black box

Het algoritme is een black box. De sollicitant weet niet hoe dit de kans op een baan beïnvloedt.

Het Financieele Dagblad.

Privacy is niet het enige mogelijke slachtoffer. Gelijke behandeling is het andere. Zoals bij meer sollicitatiebots, worden ook hier mensen in hokjes geplaatst. De programmeurs weten niet hoe het oordeel tot stand komt. Hun algoritme is een 'black box'. De sollicitant weet niet hoe dit de kans op een baan beïnvloedt.

De Raad van Europa, de officiële waakhond voor de mensenrechten, publiceert nu een rapport dat waarschuwt voor nieuwe soorten discriminatie door computers. Niet alleen de illegale, waar huidskleur, geslacht of geloof een rol spelen, maar ook de subtiele, waarbij de computer 'nieuw uitgevonden klassen' van mensen introduceert.

Het rapport vraagt ook om nieuwe regels. Want 'slimme' systemen rukken op bij politie, bedrijven, banken en overheid, en de wetgeving laat te wensen over.

De auteur van het rapport is Frederik Zuiderveen Borgesius, hoogleraar computerrecht in Amsterdam en Nijmegen. Het gevaar, volgens de jurist, zit meestal niet in het algoritme zelf, maar in de selectie van gegevens waarmee het wordt 'getraind'. Zit dáár een vooroordeel ingebakken, dan wordt dat steeds herhaald door het algoritme.

'Het meest verontrustend', zegt Zuiderveen Borgesius, 'is dat de meeste organisaties goede bedoelingen hebben. Ze willen niemand discrimineren. Maar ze hebben niet door dat hun systemen discriminerende effecten hebben.'

➔ **Vervolg op pagina 3**

→ Vervolg van pagina 1

Een van de bekendste voorbeelden is de 'sollicitatierobot' van Amazon. Het Amerikaanse bedrijf wilde de dagelijkse lawine aan cv's hanteerbaar maken, en experimenteerde in het geheim met een algoritme dat geschikte kandidaten selecteert. Eind vorig jaar ging de stekker eruit, omdat het systeem een onuitroeibare voorkeur voor mannen had.

Het probleem zat in de trainingsdata: dat waren de cv's die het bedrijf sinds 2004 ontving. In die periode solliciteerden vooral mannen. Daardoor had het systeem zichzelf geleerd dat vrouwen ongeschikt zijn. Een tussentijdse ingreep om de afwijking te fixen, werd door het systeem zelf omzeild: het selecteerde op woorden die mannen vaak gebruiken ('execute', 'capture').

Minder bekend, aldus Zuiderveen Borgesius, zijn de 'grote effecten' die 'heel veel kleine AI-beslissingen samen' kunnen hebben. Zoals prijsdifferentiatie door algoritmes, waarbij de computer beslist welke klant welke prijs krijgt. Dat kan ertoe leiden 'dat bepaalde groepen in de samenleving consistent meer betalen. Zulke differentiatie versterkt sociale ongelijkheid.'

Zijn rapport waarschuwt voor dit nieuwe type discriminatie, 'omdat die zich onttrekt aan huidige wetgeving'. Daarom stelt hij verschillende soorten regels voor. 'We hebben niets aan één regel voor alle algoritmes', zegt hij. 'Soms is bestaande wetgeving genoeg, maar bij grote risico's voor discriminatie moet je extra regels invoeren.'

Zijn voorbeelden komen vrijwel allemaal uit de VS. De auteur zelf noemt dat 'opvallend', maar heeft geen duidelijke verklaring. 'Het komt deels door de taalbarrière binnen Europa', zegt Zuiderveen Borgesius. 'Omdat ik Engels lees en geen Roemeens, weet ik eerder wat er in de VS gebeurt dan in Roemenië. Maar misschien hebben we in Europa ook te weinig aandacht voor discriminatie.'

Algorithm Watch, een particulier netwerk van onderzoekers, bracht vorige maand in kaart hoe AI-beslissingen oprukken in Europa. In veel gevallen gaat het om mogelijke discriminatie. De desbetreffende organisaties laten niemand rondkijken in de ingewanden van hun systeem.



Frederik Zuiderveen Borgesius

FOTO: DIRK GILLISSEN



'De meeste organisaties hebben niet door dat hun systemen discriminerende effecten hebben'

De Britse politie, bijvoorbeeld, ligt nu onder vuur van mensenrechtenorganisaties. De actiegroep Liberty vroeg alle politiekorpsen of ze algoritmes gebruiken om misdaad te voorspellen. Veertien korpsen zetten slimme computers in om te bepalen waar ze vaker een rondje door de buurt rijden. Eén korps gebruikt ze voor de beslisning of een verdachte wel of niet in voorlopige hechtenis blijft. De data zijn een mengsel van oude politiekegevens, postcode en leeftijd.

In Nederland is vooral SyRI omstreden: het systeem dat fraudeurs en andere criminelen moet opsporen door de koppeling van overheidsgegevens. Het ministerie van Sociale Zaken wil het algoritme niet prijsgeven, omdat het daarmee nutteloos zou worden. In de loop van dit jaar moet de staat zich verantwoorden, in een rechtszaak die is aangespannen door verschillende organisaties voor mensenrechten.