

**Arjan van Hessen**

Onderzoeker taal- en spraaktechnologie aan de Universiteit Twente

**ND.NL/OPINIE**

**BEELD ISTOCK**

# Strop dreigt voor IT-sector

HET IS GOED DAT PERSOONSgebonden data worden beschermd, om misbruik te voorkomen. Maar het is jammer als door te veel beperkingen de Europese industrie de strijd verliest van Amerikaanse of Chinese concurrenten.

In mei treedt de nieuwe Europese wetgeving voor databescherming, de General Data Protection Regulation (GDPR), in werking. Hierdoor wordt het voor bedrijven een stuk lastiger om structureel persoonsgebonden data te verzamelen, langdurig te bewaren en met anderen te delen.

Dat is een strop om de nek van Europese techbedrijven. Zij hebben deze data hard nodig in hun strijd om het nieuwe IT-goud: het taalvermogen.

De huidige toepassingen van kunstmatige intelligentie, zoals zelfsturende auto's of 'intelligent' communicerende chatbots (een geautomatiseerde gesprekspartner), duwen de mens naar de achtergrond. Mensen blijven dikwijls wel eindverantwoordelijk, maar het leeuwendeel van het werk wordt door software gedaan. Die software kan patronen in grote hoeveelheden ruizige data ontdekken. Maar échte, sterke kunstmatige intelligentie vraagt meer.

## taal

Wat ons mensen bijzonder maakt, is de mogelijkheid tot denken in concepten, in taal. Alles wat je verzint, denkt en voelt, gebeurt aan de hand van taal. Daarmee bedoel ik niet het Nederlands of Italiaans, maar de interne stem waarmee we in ons hoofd met onszelf communiceren. Dit taalvermogen stelt ons in staat te overleggen en onze gedachten te verwoorden.

Om kunstmatige intelligentie op menselijk niveau te brengen, moeten computers leren 'denken in taal'. Het bedrijf dat dit het eerst voor elkaar weet te krijgen, heeft goud in handen. Daarvoor zijn grote hoeveelheden data nodig. Maar door de nieuwe Europese wet zullen die vanaf volgend jaar niet meer of slechts met grote moeite beschikbaar zijn voor Europese techbedrijven. Daarmee wordt de waarschijnlijkheid dat zij op internationaal niveau kunnen blijven meedoen met de concurrenten uit de VS en China, met de dag kleiner. Juist nu heeft onze IT-sector meer data nodig.

## leasebranche

Neem de leasebranche. Via internet staat een groeiend aantal applicaties en apparatuur met elkaar in verbinding. Dit genereert een enorme hoeveelheid data, die tot nu toe voornamelijk negatief in het nieuws komt.

Er is echter ook een – vaak onderbelichte – positieve kant. Dezelfde data kunnen gebruikt worden om bedrijfsprocessen te stroomlijnen, kosten en tijd te besparen en zelfs het milieu te ontzien, door een verlaging van energiekosten en uitstoot.

---

'Thermostaten' kunnen verkeersspecialist helpen bij inrichting wegen.

Data op individueel niveau, die tot een persoon kunnen worden herleid, kunnen gevaarlijk zijn. Maar geanonimiseerde data op groepsniveau kunnen juist veel voordelen opleveren.

## **slimme thermostaat**

Met de data van een 'slimme thermostaat' zou je kunnen uitvoeren hoe laat iemand meestal zijn of haar huis verlaat en weer terugkomt. Dat kan het inbrekersgilde op verkeerde ideeën brengen.

Toch zijn dit soort persoonlijke data in combinatie met talige data goud waard. Wanneer een (stem)commando aan de thermostaat wordt gegeven, bijvoorbeeld 'ik ga over een halfuur weg', dan weet de slimme thermostaat dat de temperatuur omlaag kan.

Met de geanonimiseerde data van een heleboel van deze thermostaten kunnen verkeersspecialisten helpen bij een optimale inrichting van wegen en het openbaar vervoer. Zij kunnen snel zien waar en wanneer pieken en dalen kunnen optreden. Gemeenten zouden juist dit soort gegevens beschikbaar moeten maken, in de hoop dat bedrijven er handige diensten mee ontwikkelen – diensten waarvan we nu nog niet weten dat we ze nodig hebben.

## **schijnveiligheid**

Een wet als de GDPR stelt terecht strenge eisen aan de beheerders van data die tot personen te herleiden zijn. Maar zulke wetten zorgen voor schijnveiligheid; de data zijn er namelijk – en kunnen misbruikt worden. Het enige wat we kunnen doen, is het juist zo inrichten dat je de kans op misbruik zo klein mogelijk houdt. Bijvoorbeeld door 'safe rooms': ruimten waar data zonder internetverbinding worden opgeslagen, zodat ze niet gestolen of misbruikt kunnen worden. Maar door op safe te spelen en te besluiten dat dit soort data helemaal niet meer gebruikt mogen worden, wordt het kind met het badwater weggegooid. De Europese IT-sector heeft toegang tot dit soort data nodig, wil hij de aansluiting met de rest van de wereld niet missen.

Begrijp me niet verkeerd: GDPR is een uitstekende stap op weg naar meer en betere bescherming. Maar het wordt erg lastig gemaakt

om data die tot personen te herleiden zijn, te gebruiken voor het trainen van software.

Wat nodig is, is een constructie waarbij én de GDPR wordt gerespecteerd én de waardevolle, persoonlijke data toch gebruikt kunnen worden voor ontwikkelingen in de techsector. Alleen dan hebben Nederlandse en Europese IT-bedrijven een kans in de zoektocht naar de heilige graal van de IT: de omgang met de talige mens. <

# Vrij Nederland

€ 0,89



207

14 min lezen

## Zo gaat het nu met de 'verloren generatie': de pas afgestudeerden die door de crisis geen werk vonden

Oké, de werkloosheid onder jongeren is gedaald van 15 procent op het hoogtepunt naar 9 procent. Maar dat betekent niet dat alles nu koek en ei is.

# Gezondheid

€ 0,19



693

4 min lezen

**D**

€ 0,39



55

3 min lezen

**De directeur van**