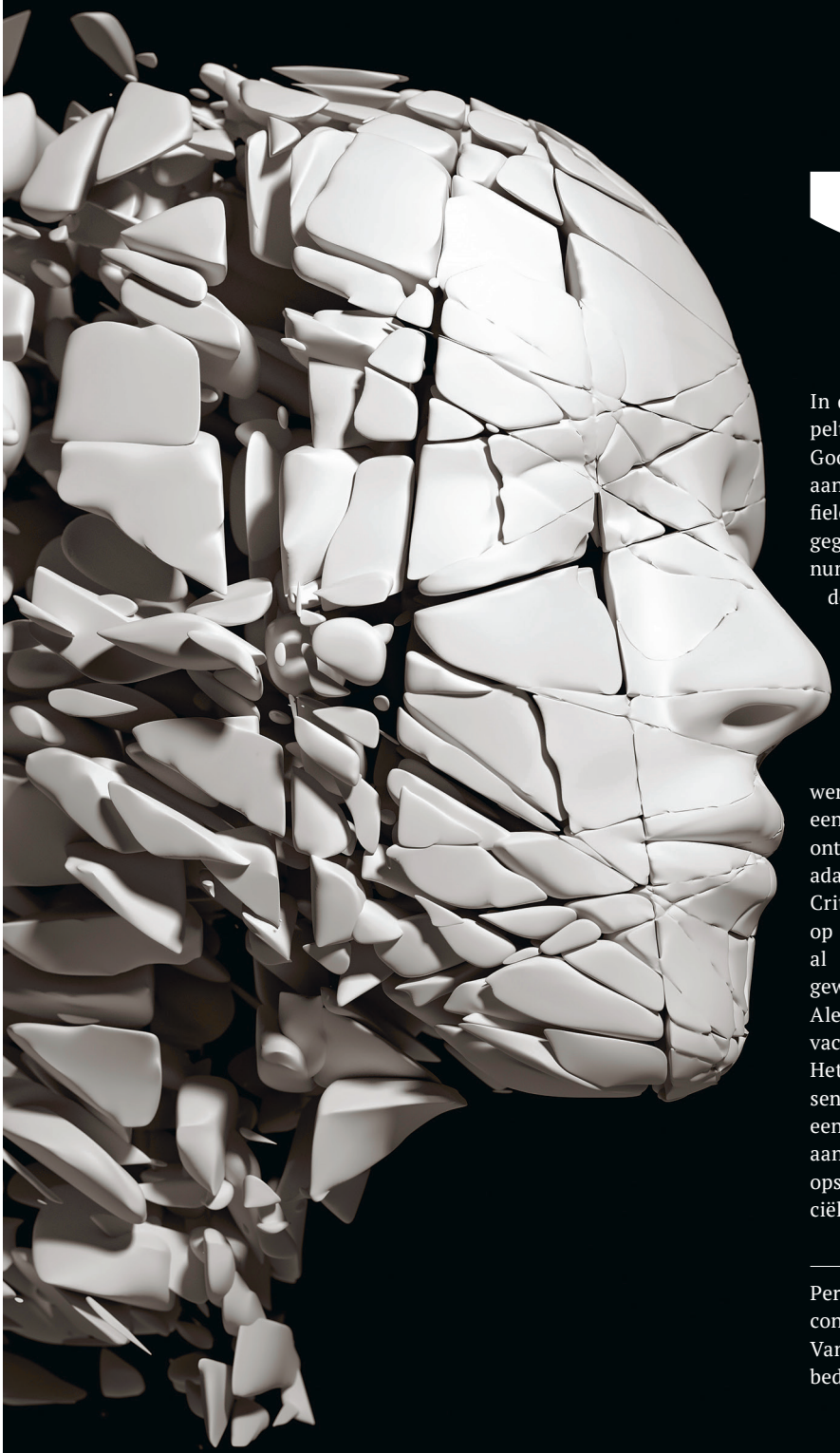
TUSSEN DROOM EN DATA

Noem data de nieuwe olie, als je er verkeerd mee omgaat, krijg je vuile handen. Kijk uit voor datalekken, privacybezwaren en bevooroordeelde artificial intelligence.

Tekst **JAN BLETZ** Beeld **GETTY**



In de dystopische roman *De Cirkel* koppelt het gelijknamige bedrijf (officieus Google) steeds meer data van mensen aan hun Cirkel-account. Niet alleen profielen en e-mails maar ook gezondheidsgegevens, creditcard- en burgerservicenummers kunnen zo worden opgeslagen door De Cirkel. Bovendien plaatst het bedrijf over de hele wereld camera's, zodat niets zich aan het alziend oog van De Cirkel kan onttrekken. Het einddoel is de zogenoemde voltooiing van De Cirkel, oftewel de vorming van een gesloten netwerk dat alle mensen verbindt. Zo kan een wereld van volledige transparantie ontstaan, waarin 'Privacy is diefstal' het adagium is.

Critici van Dave Eggers' boek merken op dat de toekomst in deze roman al grotendeels werkelijkheid is geworden: van Nest videocamera's tot de Alexa-microfoons: we geven onze privacy al volledig over aan techgiganten. Het boek doet daardoor niet erg verrassend aan. Eerder onderstreept het nog eens de grenzen waartegen ondernemers aanlopen naarmate ze steeds meer data opslaan en willen inzetten voor commerciële doeleinden.

———PRIVACY

Persoonsgegevens en andere data rond consumenten zijn geld waard, veel geld. Vandaar dat Google, Facebook en andere bedrijven ervoor betalen, of liever ge-

→

zegd: dat ze hun miljardenbusiness hebben gebouwd op de handel in data. Die krijgen ze binnen met 'gratis' diensten waarvoor consumenten betalen door hun hebben en houden prijs te geven, van waar ze zich bevinden tot hun interessen. In het verleden gebeurde dit zonder dat consumenten het wisten, maar de privacyregels zijn voortdurend in ontwikkeling en binden bedrijven, overheden en andere organisaties aan steeds striktere voorwaarden.

Organisaties in Europa zijn verplicht data van privépersonen goed te beheren en te beschermen volgens de General Data Protection Regulation (GDPR). De wet bevat privacyregels die gelden voor alle ondernemingen die in Europa actief zijn, dus ook Amerikaanse bedrijven als Google en Facebook. Bedrijven die zich er niet aan houden, lopen tegen boetes aan tot 4 procent van hun jaaromzet of 20 miljoen euro per overtreding.

'Ik zou nooit betalen voor ransomware'

"Voor veel bedrijven betekent de wet dat ze aan de slag moeten", zegt ICT-jurist Arnoud Engelfriet, auteur van boeken als *De wet op internet*. Bedrijven moeten volgens de GDPR een duidelijk beeld hebben van waar en hoe zij persoonsgegevens verwerken. Dat kan een hoop administratieve rompslomp opleveren: van alle data moet in kaart worden gebracht waar ze vandaan komen, hoe ze verwerkt worden, hoe ze beveiligd zijn en wie er toegang toe heeft. Engelfriet wijst erop dat bedrijven bijvoorbeeld in duidelijke richtlijnen moeten vastleggen wanneer ze informatie vernietigen die sollicitanten hebben toegestuurd: cv's mogen vier weken bewaard blijven tot maximaal één jaar, mits de sollicitant daar expliciet toestemming voor geeft. "Voor de meeste bedrijven bepaald geen business as usual."

De grondgedachte in de GDPR is een goede bescherming van de privacy van burgers en consumenten. Bedrijven die hun klanten willen benaderen, moeten voorafgaand aan iedere actie voortaan eerst een analyse uitvoeren van de impact op hun privacy en klanten van tevoren toestemming vragen om hen te mogen benaderen. "Expliciete toestemming, welteverstaan. Bedrijven mogen dus niet een vooraf aangevinkt vakje verstoppen in hun algemene voorwaarden, zodat een klant ermee akkoord gaat dat hij een nieuwsbrief krijgt toegestuurd, zoals nu nog vaak gebeurt. De klant krijgt bovendien het 'recht van verzet': de toestemming intrekken om persoonlijke informatie te gebruiken moet net zo makkelijk worden als toestemming geven."

De GDPR leest als een flinke lijst verplichtingen voor ondernemers. Consumenten hebben het recht om 'vergeten' te worden, waarbij alle persoonsgegevens moeten worden gewist door de partij die daarop wordt aangesproken. Dataportabiliteit moet het overstappen naar een concurrent, met medeneming van alle privégegevens, versoepelen en consumenten moeten makkelijker inzage krijgen in de data die over hen zijn opgeslagen.

Bedrijven die grootschalig persoonsgegevens verwerken, moeten een 'functionaris voor de gegevensbescherming' (FG) aanstellen, die onder andere verantwoordelijk is voor de melding van datalekken aan de Autoriteit Persoonsgegevens. Wat 'grootschalig' exact inhoudt, zegt de wet niet expliciet, maar te denken valt aan verzekeraars, energiebedrijven of marktonderzoeksbureaus.

Datalekken waarbij persoonsgegevens zijn verloren of gestolen, moeten sneller worden gemeld dan onder de vorige Wet Bescherming Persoonsgegevens en bedrijven lopen eerder tegen een boete op: in de WBP moest nog sprake zijn van opzet of grove schuld, deze bepaling verdwijnt in de GDPR. Engelfriet vindt de wet een goede zaak. "Het recht op privacy is een grondrecht, en ik kan alleen maar blij zijn dat hierop wordt toegezien." De wet brengt voor bedrijven ook meer mee dan alleen een lastenverzwaring. "Het betekent in elk geval dat veel 'shady' bedrijven die handelen in e-mailadressen de voet wordt dwars gezet: adressen verzamelen via quizjes op

‘De GDPR zet shady bedrijven de voet dwars’

internet zonder dat duidelijk is wat daar vervolgens mee gebeurt, kan straks niet meer. Voor bedrijven die netjes zaken doen is dat heugelijk nieuws.”

———CYBERSECURITY

Een goede beveiliging van data is niet alleen een wettelijke plicht. Cybersecurity is ook vanuit bedrijfseconomisch oogpunt van belang. Complete bedrijfsprocessen kunnen stilvallen door een cyberaanval, denk aan de Petya-ransomware die in de vroege zomer van 2017 de containerterminal van Maersk in Rotterdam lamlegde. Als gevolg van de onbruikbare IT-systemen konden er dagenlang geen schepen gelost of geladen worden op de terminals van zusterbedrijf APM in de Rotterdamse haven. Daar werd uiteindelijk met nieuwe apparatuur een compleet nieuw netwerk opgezet, omdat niet alle computers konden worden gerepareerd. Schade: tussen de 250 en 300 miljoen euro.

Door aanvallen van hackers kan ook gevoelige informatie op straat komen te liggen. Denk aan de gegevens over vreemdgangers die hackers wisten te stelen van de Amerikaanse datingsite Ashley Madison. Het kostte de site alleen al 10 miljoen euro aan schikkingen met gedupeerde gebruikers. De reputatie van je bedrijf kan zomaar te grabbel worden gegooid, denk aan de hack van het e-mailsysteem van consultantsbureau Deloitte, dat toch uitgerekend zelf - onder veel meer - cybersecurityadviezen geeft. Alleen al bij de grootste Nederlandse ondernemingen en overheden leidt cybercriminaliteit tot een schadepost van rond de 10 miljard euro per jaar.

Volgens Michiel Prins, mede-oprichter van HackerOne, een platform dat ‘ethische hackers’ beloont als ze bedrijven at-

→

tenderen op kwetsbaarheden in hun systemen, moeten ondernemers vooral uitkijken voor ransomware, software die je data versleutelt en enkel tegen betaling (dikwijls middels het moeilijk te traceerbare Bitcoin) weer vrijgeeft.
Ransomware is betrekkelijk makkelijk te bestrijden. "De schade is wel aanzienlijk: dit jaar hebben bedrijven naar schatting al 1 miljard dollar betaald aan 'losgeld'", aldus Prins. Maar dat had makkelijk voorkomen kunnen worden: als je software up-to-date is en je hebt een goede virusscanner en firewall, loop je weinig risico. En als je goede backups hebt, kun je de schade snel herstellen als je toch besmet bent geraakt. "Ik zou in elk geval nooit betalen: je hebt geen garantie dat degene die je software heeft versleuteld die weer vrijgeeft. Je loopt bovendien kans dat je op een 'suckers list' terechtkomt met andere gewillige slachtoffers."
Een veel groter gevaar is volgens Prins dat veel bedrijven hun producten - alles van industriële installaties en infrastructuur tot ijskasten en waterkokers – aan internet koppelen. De software die de bedrijven ontwikkelen is nogal eens onveilig en kwetsbaar voor allerlei hacks. Bedrijven die een hack willen voorkomen, kunnen het beste een apart netwerk aanleggen voor hun connected apparaten. "En bedrijven die zelf Internet of Things-apparaten aanbieden, kunnen het beste een regeling voor *responsible disclosure* invoeren, zodat hun afnemers problemen makkelijk kunnen aanmelden. De Nederlandse overheid raadt dit overigens ook aan." Een leidraad voor een responsible disclosure kan worden gedownload bij het Nationaal Cyber Security Centrum.

BETROUWBAAR

De ondernemer die zich netjes aan de wet houdt en zijn of haar data goed beveiligt zal ook moeten investeren in de kwaliteit van die data. De data moet betrouwbaar zijn ofwel 'integer', en bruikbaar. Ook zullen de modellen waarmee je analyse werkt, moeten deugen.
Aan de data-integriteit schort nogal eens het een en ander, zegt Anneke Wieling, managing director van Protiviti Nederland, adviesbureau op het gebied van riskmanagement, business consulting en audit. "Het komt erop neer dat data zo lang mogelijk relevant blijven en voor iedereen hetzelfde betekenen. Daarvoor neem je beheersmaatregelen: wie heeft toegang tot de data? Wie mag de data aanpassen? Op welke tijdstippen wordt gekeken of de data nog wel relevant zijn? Wanneer wordt gecontroleerd op doublures?"
In de praktijk worden klantdata vaak niet goed bijgehouden; het klantenbestand wordt elke keer dat iemand verhuist of overlijdt minder bruikbaar. Ook een misser: informatie over het koopgedrag van klanten die niet kan worden gekoppeld aan de adresgegevens in het CRM-systeem. Of er worden allerlei systemen op elkaar gestapeld zonder dat deze goed aansluiten. Summiere informatie die afkomstig is uit het ene systeem moet worden overgetikt om het verder te verwerken, met alle risico's op fouten van dien.
Zo eindig je als bedrijf met onvolledige, inconsistente, corrupte of gedupliceerde data. "En loop je het risico dat er vanuit de wel bekende data geen relevante analyses komen, of onjuiste beslissingen worden genomen. In die zin is het 'garbage in, garbage out', hoe geavanceerd de modellen die op de data worden toegepast ook zijn."

ZUIVERE MODELLEN

Veilig, schoon en correct. Maar ook als dat op orde is, is het de vraag of de data die je als ondernemer gebruikt, relevant genoeg zijn voor het doel dat je voor ogen hebt. Een risico dat steeds meer opduikt naarmate we intelligentere software gebruiken om onze data te analyseren is dat de modellen die worden gebruikt om patronen in data te herkennen, 'zelfversterkend' kunnen zijn.
Denk aan de algoritmen die sites als Facebook gebruiken om nieuws af te stemmen op het interesseprofiel van de gebruiker: het leidt er al gauw toe dat mensen hun

eigen werkelijkheid scheppen in de wereld van sociale media, waarin alleen boodschappen van gelijkgezinden doordringen tot hun 'bubble'.
Elders gebruiken online aanbieders soortgelijke algoritmes om klanten aan zich te binden: de befaamde recommendation engines, die persoonlijke aanbevelingen doen op basis van iemands gedrag uit het verleden: waar is naar gezocht, wat is er al aangeschaft? Aangenomen wordt dat iemand die interesse heeft in een bepaald type product of bepaalde content wellicht interesse heeft in meer van hetzelfde, of wil hebben wat vergelijkbare gebruikers kochten. Hier is de aanname dat gebruikers met dezelfde voorkeuren tot over de hele linie een zelfde smaak hebben, en dat deze in de tijd niet zal veranderen. Een bubble die aanbieders omzet kan kosten.

'Bij data is het garbage in, garbage out'

UITKIJKEN MET AI

Helemaal uitkijken wordt het als we in de toekomst vaker zelflerende systemen loslaten op data. Dan ontstaat het gevaar dat inmiddels *algorithmic bias* is gedoopt: de modellen waarmee de kunstmatig intelligente systemen werken, zouden met eenzijdige resultaten naar boven kunnen komen.
Op basis van data-analyse wordt nu al besloten wie een lening krijgt en wie niet, wie wordt uitgenodigd op een sollicitatiegesprek en wie niet en in Amerika zelfs wie vervroegd in vrijheid wordt gesteld. Als we steeds belangrijkere beslissingen overlaten aan zelflerende systemen, lopen met name minderheden en minder draagkrachtigen eerder het risico buiten de boot te vallen.
In de VS onderzoekt het AI Now Institute, opgezet door de universiteit van New York, de sociale implicaties van kunstmatige intelligentie ofwel artificial intelligence (AI). Wie met AI werkt, zo dringt AI Now aan, zal in de gaten moeten houden of het model komt met uitkomsten die passen bij wat maatschappelijk verantwoord is. Het instituut roept overheidsinstanties op belangrijke beslissingen niet meer over te laten aan 'black box AI' zolang de systemen hun oordeel niet kunnen uitleggen en daar ook geen verantwoording over kunnen afleggen.
Ook bedrijven staan vanzelfsprekend voor de uitdaging hun zelflerende algoritmes bij de les te houden – waarbij het de vraag is of elke black box nog wel te openen is om uit te vinden hoe de analyse plaatsvindt.
In de techwereld tekent zich wat dat betreft een stammenstrijd af: een kamp met daarin onder anderen Elon Musk waarschuwt al jaren dat artificial intelligence zijn eigen leven gaat leiden en uiteindelijk zal ontaarden in levensgevaarlijke robots. Steve Wozniak, mede-oprichter van Apple, ziet wel een vreedzame toekomst voor zich – waarin kunstmatig intelligente systemen mensen als huisdieren houden. Het tegenkamp, aangevoerd door Mark Zuckerberg, verwijt dergelijke 'doemzeggers' dat ze een te beperkt begrip hebben van de materie en onverantwoord handelen met hun bangmakerij. Maar uitgerekend in zijn eigen Facebook Artificial Intelligence Researchlab ging de stekker uit een experiment waarin twee spraakrobots in een onderlinge dialoog een onbegrijpelijke eigen taal bleken te ontwikkelen. Het experiment schoot zijn doel voorbij, was de offficiële lezing.
Maar als zelfs een briljante wetenschapper Stephen Hawking spreekt van 'het ergste wat onze beschaving is overkomen', telt een gewaarschuwd ondernemer voor twee. Data is de nieuwe olie, maar zorg dat je er niet mee knoeit.