



TEKST: MALOU VAN HINTUM, BEELD: JELLE HOOGENDAM

Het is geen hype, zeggen ze eensgezind. Integendeel, het belang van 'big data' zal alleen maar groter worden. Joost Kok (hoogleraar Informatica en hoogleraar Geneeskunde) en Gerrit-Jan Zwenne (hoogleraar Recht en de informatiemaatschappij) kruisen de degens. Kok ziet grootse mogelijkheden, Zwenne vreest kafkaïaanse toestanden.

Joost Kok, hoogleraar Informatica en hoogleraar Geneeskunde (**JK**): 'Er is geen getal om aan te geven wanneer er sprake is van big data en wanneer niet. Het verschil tussen grote dataverzamelingen en big data is dat bij big data de standaardmethoden niet meer werken. De statistiek stort in, we gaan benaderende methoden gebruiken in plaats van exacte. Het is vechten tegen een soort grens. Sinds een jaar of vijf hebben we de computerkracht om dat ook echt te doen.'

Gerrit-Jan Zwenne, hoogleraar Recht en de Informatica-maatschappij (**GZ**): 'Terwijl wetenschappers vroeger met steekproeven werkten, kunnen ze nu met de volledige dataset werken.'

JK: 'Daardoor kun je 'dommere' methoden gebruiken. Tot nu toe maakte je een prachtig model en liet je dat op je data los. Nu kun je op grond van de data zelf al dingen zeggen. Aan twitterberichten kun je sneller zien waar een aardbeving is dan aan seismografische meters, aan het aankoopgedrag van mensen zie je sneller hoe de griep zich verspreidt dan aan de cijfers van huisartsen. Dan heb je het over correlaties, maar daar kun je heel veel mee. Nadenken over correlaties kan leiden tot creatieve vondsten.'

GZ: 'Viktor Mayer-Schönberger, hoogleraar Internet governance aan de universiteit van Oxford, vertelde dat een ziekenhuis op te vroeg geboren baby's heel veel sensoren had geplakt en alle mogelijke data had verzameld. Hierdoor werd ontdekt dat als zo'n kindje 24 uur heel stabiel is, er een moment komt waarop er iets misgaat - terwijl iedereen juist dacht dat het goed was als een baby lang stabiel was. Op basis daarvan is de behandeling aangepast. Dat is positief.'

'Maar we moeten tegelijkertijd ook waakzaam zijn voor mogelijk misbruik. Ik las op Ouders Online een verhaal over drie kinderen die uit huis waren geplaatst op grond van niets anders dan een risicoprofiel: u woont in een buurt die niet zo goed is, u verdient niets, u bent laag opgeleid, er is een hennepplantage in uw straat, enzovoorts. Na 'Savanna' valt het niet mee hulpverleners die met te weinig middelen gruwelijke ongelukken moeten tegengaan, ervan af te houden om technieken als profiling toe te passen. Technieken die voorbij gaan aan de onschuldpresumptie (het uitgangspunt dat iemand onschuldig wordt geacht totdat het tegendeel is bewezen), en die het aanzienlijke risico van stigmatisering en discrim-

inatie hebben. En van ontmenselijking: computer says no. Het is Kafka. Bij privacy denken we vaak aan Orwell's Big Brother, bij big data gaat het vaak meer om Kafka. Er worden beslissingen genomen die je niet kunt herleiden en waar niemand zich verantwoordelijk voor acht; het is de computer die zegt 'u bent een probleemgeval.'

JK: 'Je hebt altijd extra informatie nodig om tot een profiel te komen, dat kan nooit alleen op basis van computerinformatie. Er zijn altijd mensen nodig om zo'n stap te zetten. Ik zie data als een grondstof waarmee je van alles kunt doen, veel meer dan nu gebeurt. Er zijn genoeg positieve voorbeelden. DNA-analyses waardoor moordenaars worden gevonden, of waardoor we binnen een paar uur iedereen kunnen identificeren die in een verongelukt vliegtuig zat. Door big data kun je dingen met elkaar vergelijken: waarom is het ene ziekenhuis efficiënter dan het andere? Waar kunnen ziekenhuizen het beste staan, hoe kunnen ambulances het beste rijden? In Leiden kun je op grond van telefoonsignalen precies zien hoe mensen door de stad bewegen. Zo kun je zien of de wegomleggingen wel goed functioneren. Op dat niveau kun je zo'n stad enorm verbeteren.'

GZ: 'Een aanbieder van autonavigatiediensten kreeg ooit het verzoek van een gemeente om een bepaald dijkje uit de route te halen zodat er geen verkeer meer over zou komen. Wat ik weet is dat deze aanbieder dat niet heeft gedaan: als je niet wilt dat iemand er overheen rijdt, zet je er maar een verkeersbord neer. Stel dat een wethouder de beschikking krijgt over dit soort instrumenten, zodat hij niet meer met behulp van verkeersborden en verordeningen stuurt hoe het verkeer door de stad gaat, maar door de stoplichten wat langer aan of uit te zetten met intelligente algoritmes waardoor hij de doorstroming weet te vergroten... Daarop moet controle zijn. Want wat als hij die techniek gaat gebruiken om demonstraties tegen te gaan?'

JK: 'Vervang demonstraties door de uitbraak van een ziekte, ebola bijvoorbeeld, en je krijgt dankzij dit type techniek de ziekte onder controle. Of gebruik het om stropers te stoppen door allerlei controlesystemen op elkaar af te stemmen. Dan is het weer goed!'

GZ: 'Er kan iets veranderen in de machtsverhouding tussen burger en overheid. Voor een jurist moet daar iets tegenover

Leiden Centre of Data Science

Afgelopen najaar heeft prins Constantijn in Leiden het Leiden Centre of Data Science (LCDS) geopend. De faculteit Wiskunde en Natuurwetenschappen vormt samen met het LUMC (Leids Universitair Medisch Centrum) de kern van het LCDS, dat ook als

expertisecentrum fungeert binnen de Universiteit Leiden, voor het Bio Science Park Leiden en voor het regionale ruimtevaartcluster. Het data-science-onderzoek in Leiden onderscheidt zich van de rest van Nederland doordat het fundamenteeler is en een grote inbreng van statistici heeft. Ook ligt de focus op bio-data.

Onderzoek richt zich op uiteenlopende onderwerpen als het supersnel analyseren van ziekten, een groter begrip van de geschiedenis van het heelal, het opsporen van fraude bij zorgdeclaraties, het spelgedrag van kinderen op het schoolplein en het beter laten presteren van topsporters.

Het Centrum voor Recht en Digitale Technologie eLaw onderzoekt hoe het recht kan bijdragen aan een goede werking van informatie- en communicatietechnologie, en het gebruik ervan door burgers, bedrijven en overheden. Ook ontwikkelt het nieuwe modellen en instrumenten voor regulering.



Joost Kok:

'Ik zie data als een grondstof waarmee je van alles kunt doen, veel meer dan nu gebeurt'

Gerrit-Jan Zwenne:

'Bij privacy denken we vaak aan Orwell's Big Brother, bij big data gaat het vaak meer om Kafka'

FOTO: ROB OVERMEER

staan. *Balancing powers, trias politica, level playing field*. Iets anders is het idee 'we gaan verzamelen en bedenken later wel wat we ermee doen'. Zo begon Google met Streetview. Maar dataprotectiewetgeving schrijft voor dat je van tevoren een welbepaald uitdrukkelijk omschreven rechtvaardig verzameldoel bepaalt, dat je meedeelt aan de datasubjecten. In dit voorbeeld leek Google daaraan voorbij te gaan.'

JK: 'Wij doen samen met Philips en het LUMC onderzoek naar veroudering. Daarvoor plakken we mensen vol sensoren en proberen we continu aan ze te meten. We meten ontzettend veel, maar hoe halen we eruit wat we willen weten? Omdat dat allemaal nieuw is, hebben we dat moeten leren. Dat betekent dat je, inderdaad, gewoon maar begint met enorme hoeveelheden data te verzamelen.'

GZ: 'Er is bij big data het risico van *function creep*; dat je telkens een op zichzelf nuttige functionaliteit erbij bedenkt - denk bijvoorbeeld aan de ov-chipkaart, die steeds meer functies krijgt - en dan ergens eindigt waar je misschien niet had willen eindigen. Het is heel moeilijk om dat tegen te houden, ook omdat aan opslagcapaciteit nauwelijks meer kosten zijn verbonden. Er is de neiging gegevens hoe dan ook op te slaan, omdat er later misschien iets nuttigs mee kan worden gedaan.'

'Er zijn daarom randvoorwaarden nodig voor een acceptabele, mensvriendelijke manier van gegevensgebruik. Daarbij denk ik niet in de eerste plaats aan toestemming

- niemand leest de vijftig pagina's voorwaarden waarmee je akkoord moet gaan om een nieuwe app op je smartphone te installeren. We moeten op andere manieren verantwoordelijkheid of '*accountability*' organiseren. Mayer-Schönberger noemt in dit verband de 'algoritmisten'. Zoals accountants boekhoudingen begrijpen en kunnen beoordelen of ze kloppen of niet, zo zal deze specialistische beroepsgroep kunnen kijken of we het juiste doen met de resultaten van big data. Daarnaast kan worden gedacht aan transparantievereisten en misschien ook toezicht. Waar het uiteindelijk om gaat, is dat individuen hun vrijheid blijven houden; dat je je wil kunt uiten en dat daar iets mee wordt gedaan. Dat zijn basiswaarden waarvan we hebben bedacht dat die fundamenteel zijn, universeel zelfs.'

JK: 'Het probleem is dat we nu te veel achter de technologische ontwikkelingen aanlopen. Ik denk dat voor een aantal van die problemen technische oplossingen zullen komen, waardoor mensen kunnen bepalen wat ze aan wie ter beschikking stellen, en voor welke doelen.'

GZ: 'Dat denk ik ook. Als mensen de behoefte hebben hun data te beschermen, en daarvoor willen betalen, zal een commerciële aanbieder daarin willen voorzien en dat in zijn applicaties inbouwen.'

JK: 'En bedenk ook: er zullen altijd mensen zijn die juist graag willen dat hun data geanalyseerd worden!' □