

Bedankt voor het downloaden van dit artikel. De artikelen uit de (online)tijdschriften van Uitgeverij Boom zijn auteursrechtelijk beschermd. U kunt er natuurlijk uit citeren (voorzien van een bronvermelding) maar voor reproductie in welke vorm dan ook moet toestemming aan de uitgever worden gevraagd.

Boom

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch door fotokopieën, opnamen of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikelen 16h t/m 16m Auteurswet 1912 jo. Besluit van 27 november 2002, Stb 575, dient men de daarvoor wettelijk verschuldigde vergoeding te voldoen aan de Stichting Reprorecht te Hoofddorp (postbus 3060, 2130 KB, www.reprorecht.nl) of contact op te nemen met de uitgever voor het treffen van een rechtstreekse regeling in de zin van art. 16l, vijfde lid, Auteurswet 1912.

Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16, Auteurswet 1912) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten, postbus 3060, 2130 KB Hoofddorp, www.cedar.nl/pro).

No part of this book may be reproduced in any way whatsoever without the written permission of the publisher.

info@boomamsterdam.nl
www.boomuitgeversamsterdam.nl



Viktor Mayer-Schönberger & Kenneth Cukier
De Big Data-revolutie. Hoe de data-explosie al onze vragen gaat beantwoorden

Maven Publishing | 2013 | 303 pp. | € 22,00 | ISBN 9789490574901

Tekortkomingen in het Big Data-denken

door Rogier Havelaar

De auteur is werkzaam als business development manager bij een grote Nederlandse onderneming en als zelfstandig debat- en mediatrainer. Daarnaast is hij redacteur voor Christendemocraat.nl.

De auteurs van *De Big Data-revolutie. Hoe de data-explosie al onze vragen gaat beantwoorden* beschrijven hoe Big Data wetenschap, politiek en samenleving zal veranderen. Viktor Mayer-Schönberger, Big Data-wetenschapper en adviseur van grote bedrijven op het gebied van Big Data, en Kenneth Cukier, technologiejournalist van *The New York Times*, bundelen hun krachten om, volgens uitgeverij Maven Publishing, de gevolgen van Big Data-gebruik voor het eerst gestructureerd uiteen te zetten. Het boek is gemakkelijk te lezen en staat vol met aansprekende voorbeelden.

Het eerste hoofdstuk trapt af met het voorbeeld dat Google door het analyseren van zoekopdrachten beter voorspelt hoe een griep epidemie zich ontwikkelt dan de

gezondheidscentra in de Verenigde Staten – van wie de werkzaamheden voor een belangrijk deel bestaan uit het voorspellen van dergelijke epidemieën. Het gestructureerd doorzoeken van miljarden zoekopdrachten, likes en tweets, of het opslaan en analyseren van patiëntgegevens die voorheen niet opgeslagen werden, kan levens redden en verandert de klassieke manier waarop wetenschappers hun vak uitoefenen. Zij werken niet langer meer met zorgvuldig samengestelde steekproeven, maar gebruiken *alle* beschikbare data. Waar wetenschappers in het verleden veel moeite hadden met het verzamelen van data, hebben zij dankzij de Big Data-revolutie nu eenvoudig toegang tot een gigantische hoeveelheid gegevens.

Big Data wijst op ‘het vermogen van de samenleving om informatie op nieuwe manieren in te zetten voor het verkrijgen van nuttige inzichten of waardevolle goederen en diensten’ (p. 10), ‘waarbij een zo groot mogelijk deel van de complete gegevensverzameling betrokken wordt’ (p. 44). Dataficatie, ‘het in gekwantificeerde vorm brengen van een verschijnsel zodat het systematisch kan worden vastgelegd en geanalyseerd’ (p. 114),

B
O
E
K
E
N

is daar een belangrijk element in. ‘Bij Big Data is de som waardevoller dan de delen en als we meerdere gegevensverzamelingen met elkaar combineren is ook die som meer waard dan de ingrediënten ervan’ (p. 156).

* * *

Een gegevensbestand heeft naast het primaire doel waar het voor aangelegd wordt ook een optiewaarde. Zoektermen op internet lijken op het eerste gezicht waardeloos nadat het primaire doel bereikt is: iemand zoekt iets op internet op en wordt naar de juiste website verwezen. Het primaire doel is daarmee vervuld. De gegevens hebben echter ook nog een ‘optiewaarde’: wat kun je nog meer met de gegevens doen en hoe kun je daar geld mee verdienen?

Google heeft zijn taalprogramma’s verbeterd door suggesties te geven wanneer een *zoekterm* verkeerd wordt gespeld. *Bedoelde u: zoekterm?* Zodra de gebruiker op het juist gespelde woord klikt, weet Google dat met *zoekterm* dus *zoekterm* bedoeld wordt. Nog interessanter wordt het wanneer jouw ‘gegevensuitstoot’, het digitale kielzog dat je achterlaat, gebruikt wordt. Waar klikte je op? Hoelang bekijk je een pagina? Hoe laat je de muis over het scherm bewegen? Hoe hard druk je op de toetsen van het toetsenbord? Op basis van deze gegevens worden websites aangepast en interesses vastgelegd.

De auteurs stellen dat het hergebruiken van gegevens, het combineren van gegevens met andere datasets en het gebruiken van gegevensuitstoot grote economische kansen bieden. Hieruit kunnen verschillende bedrijfsmodellen ontstaan: bedrijven die data verzamelen en doorverkopen (zoals Twitter), bedrijven (veelal consultants) die de expertise hebben om data te gebruiken maar zelf niet over de data beschikken, of bedrijven

(veelal start-ups) met ideeën voor nieuwe toepassingen van Big Data.

Kortom: de Big Data-revolutie, waarin zo veel mogelijk aspecten van de wereld door dataficatie geschikt gemaakt worden voor analyse, brengt grote kansen met zich mee. Van het oplossen van spraakverwarring tot het redden van mensenlevens. De toekomst lijkt rooskleurig!

Toen ik begon met het lezen van het boek was mijn vooroordeel dat Big Data eigenlijk niets nieuws is. Zijn we immers niet al sinds mensenheugenis bezig om informatie vast te leggen en te analyseren? Met steeds snellere computers en ingewikkeldere algoritmes zijn we steeds beter in staat geweest om berekeningen te maken en voorspellingen te doen. Door het lezen van dit boek ben ik daar anders over gaan denken. De auteurs hebben mij ervan overtuigd dat Big Data een nieuw fenomeen is en dat de manier waarop wij denken en handelen zal veranderen. We zullen het steeds meer als een normaal gegeven beschouwen dat bedrijven en overheden talloze gegevens van ons opslaan om daar analyses op uit te voeren. Privacy en anonimiteit worden minder belangrijk. Daarnaast zullen we steeds minder in ‘oorzaak en gevolg’ gaan denken en steeds meer in waarschijnlijkheden op basis van correlatieberekeningen. De voorspellende kracht van deze berekeningen zal alleen maar toenemen. Het ‘waarom’ zal daardoor plaats gaan maken voor het ‘wat’. Dit zal ook gevolgen hebben voor hoe wij met kennis omgaan. Kennis en ervaring worden minder belangrijk, zolang er maar data beschikbaar is om ons te vertellen wat we moeten doen. Om een introductie te krijgen van de talloze mogelijkheden van Big Data en vanwege de heldere praktijkvoorbeelden vind ik het boek een *must read*. De auteurs

schreven het boek niet voor niets zowel voor de Big Data-leek als voor de expert.

* * *

Bij dit nieuwe fenomeen horen natuurlijk ook nieuwe uitdagingen. In mijn kritiek op het boek zet ik drie tegenstellingen neer om vanuit christendemocratisch oogpunt enkele tekortkomingen in het Big Data-denken van de auteurs aan te stippen.

In de eerste plaats de tegenstelling tussen de dichter en de koopman. Stel, we bestuderen een verliefd stel dat innig afscheid neemt op een station. Big Data zal realtime via de smarthorloges van de geliefden kunnen waarnemen dat hun hartslag versnelt en dat hun lichaamstemperatuur stijgt. Camera's op het station tonen dat hun fysieke afstand zich tot nul centimeter beperkt en registreren hoelang de zoen duurt. De vraag is wat we nu eigenlijk *zien* wanneer we al deze registraties op het stel loslaten. De koopman – symbool voor de multinationals, die het best uitgerust zijn voor Big Data-verzameling en -analyse – ziet op basis van correlaties een toegenomen waarschijnlijkheid dat de vrouw behoefte heeft aan chocolade omdat ze verdrietig is dat haar geliefde vertrokken is en dat de man behoefte heeft aan kauwgom omdat uit een andere databank bleek dat zijn vrouw gisteren knoflook gegeten heeft. Maar wat *zien* we nu werkelijk? De dichter zal tot andere waarnemingen komen dan het horloge en de camera, omdat hij anders naar de wereld kijkt – namelijk vanuit nieuwsgierigheid en verwondering. Een risico van Big Data is dat de wereld platgeslagen wordt tot nullen en enen en een correlatiegetal tussen -1 en 1 met een economisch doel voor ogen. In een dergelijke wereld is er onvoldoende ruimte voor de dichter. Hierdoor zullen we uitein-

delijk minder waarnemen dan 'meer'.

Ten tweede de tegenstelling tussen verwondering en pragmatisme. Het pragmatisme zal de verwondering naar de achtergrond duwen, waardoor er minder ruimte zal komen voor subjectieve beleving en overtuiging. Zij maken plaats voor een geloof in getallen. Het boek legt uit dat correlaties steeds meer de grondslag zullen vormen voor beslissingen. Niet meer het 'waarom' maar het 'wat' komt centraal te staan. Dit zal als gevolg hebben dat mensen die een besluit baseren op het 'waarom', zonder het 'wat' aan hun zijde te hebben, ter verantwoording geroepen zullen worden. De vraag is op grond van welke norm zij al dan niet 'schuldig' bevonden kunnen worden. Heb je onverstandig gehandeld wanneer je vanuit een bepaalde overtuiging, zoals intuïtie of geloof, een keuze maakt (*waarom*), zonder dat je het *wat* aan je zijde hebt?

Aan de Big Data-filosofie ligt de levensovertuiging ten grondslag dat (bijna) alles kenbaar is, zolang je maar genoeg gegevens voorhanden hebt. Dit zal de rol van religie en persoonlijke overtuiging verder terugdringen, en het is de vraag of dat wenselijk is. Daarom pleit ik ervoor om het 'waarom' in ere te houden en de basisaannames van Big Data (het geloof in getallen en de meetbaarheid van alle dingen) ter discussie blijven te stellen.

In de derde plaats soevereiniteit versus het systeem. Wat verschaft bedrijven het recht om zo veel mogelijk van mijn gegevens uitstoot te registreren? Waarom mag een mobiele telefoonmaatschappij haar platform, de dienst die ik afneem, gebruiken voor andere doeleinden dan waar ik die maatschappij voor betaal? Het gemakkelijke antwoord op die vraag is: omdat ik toestemming gegeven heb voor deze dataregistratie.

Maar waarom heeft het bedrijf in eerste instantie om deze toestemming gevraagd? En wat is deze toestemming waard als ik tegenwoordig op iedere website uitgebreide lijsten met een privacyverklaring moet accorderen voordat ik de website kan gebruiken? Het antwoord op deze vraag hangt af van tot welke diepte je bedrijven (en overheden) wilt laten binnendringen in de persoonlijke omgeving van consumenten.

Zonder aanvullende maatregelen zal in de toekomst normaal functioneren niet meer mogelijk zijn zonder aan dataficatie onderworpen te worden. Facebook, bijvoorbeeld, blokkeert de privéberichtenfunctie op de mobiele telefoon zolang een gebruiker de app Messenger niet installeert. En wil je Messenger gaan gebruiken, dan moet je die app toestemming geven om berichten te lezen, bestanden te bekijken, locatiegegevens op te slaan en te registreren met wie er wanneer gebeld wordt. En dit zullen steeds meer apps gaan doen. De mobiele telefoonmaatschappij, je auto, de slimme energiemeter, je hardloop-

apparatuur, je navigatiesysteem, je bank, je supermarkt – talloze bedrijven, apparaten en apps zullen toestemming afdwingen alvorens je de door hen geboden functionaliteit kunt gebruiken. Daarmee heeft de Big Data-wereld het potentiële risico een totalitair

De Big Data-wereld heeft het potentiële risico een totalitair systeem te worden

systeem te worden. Als je niet instemt met de voorwaarden, kun je de producten niet gebruiken. En zo dringt het bedrijfsleven binnen op plekken die we vanuit christendemocratisch denken nog ‘soevereine kringen’ noemen. Het is de uitdaging om deze beginselen te updaten naar het Big Data-tijdperk waarin talloze bedrijven en overheden denken dat zij alles mogen registreren wat zij kunnen, in naam van de pragmatiek, om er vervolgens naar te handelen.