

# Hoe groot wordt

De zoveelste hype of the *next big thing*? En wat kunnen accountants ermee? En controllers? De ontwikkelingen, mogelijkheden, visioenen en valkuilen rond big data. 'De *data driven* maatschappij is geen keuze meer', maar de gouden eieren laten nog even op zich wachten.

TEKST: BERT BAKKER | BEELD: DREAMSTIME, MARJA BROUWER

**V**oor KPMG was het een unicum. Voor het eerst in haar bestaan riep de organisatie afgelopen jaar een eigen investeringsfonds in het leven. Doel: stimuleren dat *big data* efficiënter in de bedrijfsvoering van ondernemingen wordt ingezet. Naar verluidt hebben KPMG-vestigingen van over de hele wereld nu zo'n honderd miljoen euro in de pot gestort - als eerste tranche.

Het nieuws dat in november 2013 wereldkundig werd gemaakt, suggereert in elk geval één ding: de ontwikkelingen rond 'big data zijn van groot strategisch belang'. Voor KPMG, dus zonder twijfel ook voor de rest van de audit- en managementadviessector.

"Inderdaad, big data staat hoog op de agenda. We weten van enkele andere big four-kantoren dat ze ook fors investeren." Aan het woord is Erik Schut. Hij is managing partner bij KPMG Advisory, heeft een IT-verleden en is nauw betrokken bij het big data-fonds KPMG Capital.

De voor dit fonds gekozen opzet zegt bovendien iets, meent Schut, over hoe KPMG denkt dat de ontwikkeling van kennis rond big data, en de commerciële toepassingen ervan, zou moeten verlopen. "Wereldwijd kunnen uit deze pot deelnemingen worden genomen in externe projecten, maar ook in interne initiatieven zal worden geïnvesteerd. Op dit moment vindt KPMG het betrokken zijn bij partijen die de verwerking en analyse van grote databestanden snel verder kunnen brengen minstens zo belangrijk als het financiële rendement."

Is big data werkelijk *The Next Big Thing*? Meer dan de zoveelste hype, opgestookt door auteurs van managementboeken?

Dat laatste zou je kunnen denken als je op YouTube een goeroe als Kenneth Cukier zijn 'evangelie' hoort verkondigen. Maar op dat zelfde kanaal vind je ook verhalen van strategieconsultants van McKinsey. Big data is volgens hen al niet meer the *next big thing*, maar *the big thing* van nu! En wie niet oplet, of niks ziet in dat analyseren van gegevens over markten, gedrag van doelgroepen, of van productie- en administratieve processen in eigen huis, die loopt nu al achter op z'n concurrent.

## Gedrag volgen

De Amerikaanse National Security Agency (NSA) tapt om terroristen en criminele organisaties op te sporen miljoenen telefoongesprekken af. Maar zelden door ze individueel af te luisteren. Belpatronen worden geanalyseerd, en als bepaalde termen of telefoonnummers vaak terugkomen wordt een signaal afgegeven. Niet mensen maar computers verwerken die verder. De grootste voorlopers in de analyse van big data en het gebruikmaken van opgemerkte patronen, zijn natuurlijk de partijen die door de aard van hun bedrijf over veel data beschikken. Zij zagen als eerste dat je eraan kunt verdienen als je de massa's data tot waardevolle informatie verwerkt. Google en Facebook. Maar ook IBM, Amazon en de meeste grote winkelketens, telecommatachappijen, banken en verzekeraars zitten er bovenop.

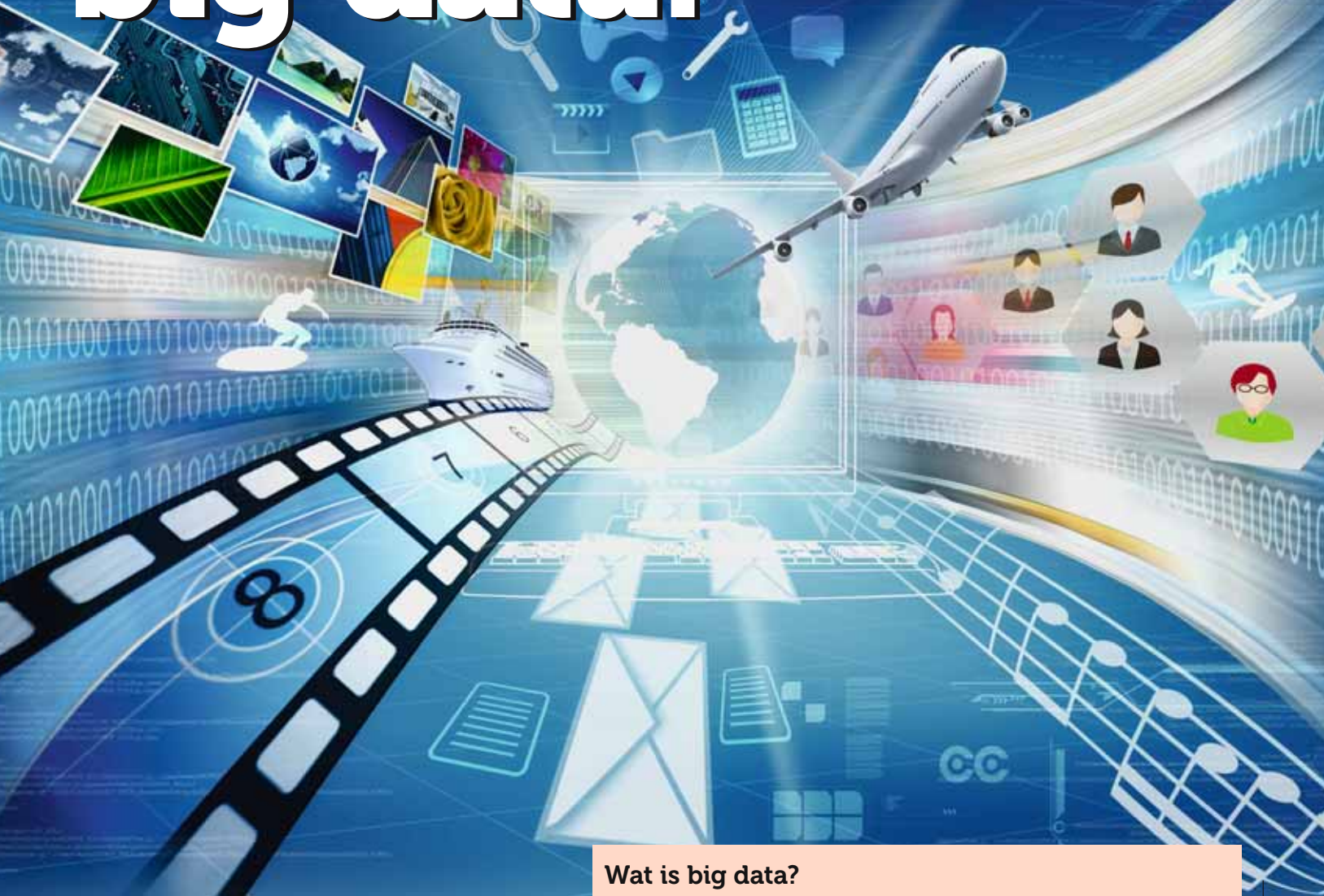
In datasets over zoekgedrag van burgers op het internet, hun belgedrag, koop- en betaalgedrag, reisgedrag, worden patronen herkend: deze dame heeft interesse in zovakanties, die meneer kijkt vaak naar Amerikaanse classic cars. Dus als je bepaalde advertenties naar hun scherm stuurt, is er een bovengemiddelde kans dat daar business uit komt. Kortom: 'handel'!

De uitdaging is om uit die exponentieel gegroeide hoeveelheid data op een efficiënte manier bruikbare info te halen met een soort digitale 'kraakinstallatie'. Dat 'kraken' houdt in: algoritmische berekeningen loslaten op datasets die als betrouwbare input zijn aangemerkt. Dat laatste is wel 'even een dingetje'. Daarover later meer in dit artikel.

Eerst die brandende vraag hoe en waar in de accountants- en adviessector, waar sowieso alles draait om

**ERIK SCHUT (KPMG): 'BIG DATA STAAT HOOG OP DE AGENDA. WE WETEN VAN ENKELE ANDERE BIG FOUR-KANTOREN DAT ZE OOK FORS INVESTEREN.'**

# big data?



informatieverwerking en -analyse, die mogelijkheden worden gezien. Wat is er gaande? En welke impact op het bedrijfsmodel valt te voorzien?

## Groot, heterogeen en exploratief

Jacques de Swart, binnen PwC verantwoordelijk voor de adviesgroep Quantitative Analysis en daarnaast deeltijdhoogleraar toegepaste wiskunde aan Nyenrode Business Universiteit, ziet het zo: "Accountants werkten natuurlijk al veel met data-analyse als ze op zoek gaan naar wat in hun ervaring uitzonderlijke boekingen zijn. En niet meer alleen terugkijkend op een boekjaar, maar ook steeds meer als *real-time* business process analytics."

Met big data kan daaraan, voorziet De Swart, iets worden toegevoegd: "Dan zoek je niet alleen maar bekende uitzonderingen gebaseerd op één dataveld maar ook naar vreemde, nog onbekende combinaties van zowel in- en als externe datavelden. Statistische technieken als multi-dimensional scaling, factor- en

## Wat is big data?

Over de definitie bestaat geen eenstemmigheid, noch over de voorbeelden.

Zie daarvoor ook dit artikel. Wikipedia formuleert het aldus:

"Men spreekt van big data wanneer men werkt met één of meer datasets die te groot zijn om met reguliere databasemanagementsystemen te onderhouden. De definitie van big data is niet altijd duidelijk. Volgens Gartner (*onderzoeks- en adviesbureau in de IT-sector, red.*) gaat het in elk geval om drie factoren: de **hoeveelheid** data, de **snelheid** waarmee de data binnenkomt en wordt opgevraagd en de **diversiteit** van de data. Daarnaast speelt ook de **mogelijkheid tot statistische analyse** een belangrijke rol: een website als YouTube verwerkt weliswaar grote hoeveelheden data maar de mogelijkheden om die data statistisch te analyseren zijn beperkt. Om die reden wordt YouTube niet altijd tot 'big data' gerekend."

clusteranalyse kunnen hierin een enorme rol in spelen, al zijn die niet altijd even bekend binnen de accountancy."

In De Swarts definitie is sprake van 'big data' als je werkt met bestanden die groot en heterogeen zijn, en waarin exploratief wordt gezocht: "Je stelt dan geen concrete vragen maar combineert datasets en kijkt of



er patronen opduiken. Je zoekt inductief naar *unknown unknowns*.”

Zo zou je, bij wijze van spreken, die befaamde *Black Swan* van Nassim Taleb op het spoor kunnen komen. Op een schaal waar kale assurance links staat en echte verbetermogelijkheden rechts, kunnen accountants door exploratief zoeken naar rechts opschuiven.

De Swart: “Dit is dé kans om te doen wat de maatschappij al zo lang van ons vraagt: niet meer sec vaststellen of de boekhouding volgens de regels is opgesteld, maar ook daarbuiten dingen zien: bedreigingen, risico’s, kansen die bestuurders zelf nog niet zagen.”

## Schaarse Early adopters

“Het vreemd dat in bedrijven, grote en kleine, de digitale data zich ophopen maar dat daarmee zo verbazend weinig gebeurt. De managementinformatie die erin verborgen zit, blijft meestal onzichtbaar en kan dus ook strategisch geen rol spelen. Doodzonde!” Het was deze observatie die Wilco Schellevis, partner bij accountantskantoor **Visser & Visser**, ertoe bracht om enkele jaren geleden een product te ontwikkelen waarmee klanten stuurinformatie kunnen halen uit data die er wel waren maar niet inzichtelijk worden gemaakt. “Het blijkt voor een bedrijf heel veel waard om beter en sneller te weten waarop je winst maakt, en waarop verlies.”

Marco Hill en Pieter de Kok, partners bij **Coney Accountants** zijn sinds 2005 bezig met het idee van *data analysis as a service*. Inmiddels is het helemaal geïntegreerd in hun adviespraktijk en de audit only assurance-praktijk. De Kok: “We helpen ondernemingen om uit eigen *company data* via *process mining* en data-analyse continu informatie te halen die van pas komt om bijvoorbeeld hun productie- of marketingprocessen te verbeteren. Dat werkt, maar wij noemen het geen big data.”

Toen big data twee jaar geleden ineens een *buzz* werd, deed Coney daaraan mee. “Maar”, zegt de Kok nu, “in 2013 besloten we te stoppen met *buzzen*. Onder de radar gaan we wel door met een paar big data-projecten die duurzaam zijn en waarvan ook de business case lijkt te kloppen. En we werken intussen verder aan SBR en gewone *small data*-toepassingen.”

## JACQUES DE SWART (PWC): ‘DIT IS DÉ KANS OM TE DOEN WAT DE MAATSCHAPPIJ AL ZO LANG VAN ONS VRAAGT.’

### Mkb

Maar de mogelijkheden kunnen verder gaan dan waar de creëren door verrijking van het eigen product met tegelijkertijd een efficiëntere dus goedkopere werkwijze. Kantoren die voortvarend aan de slag gaan met big data kunnen een hele nieuwe categorie van adviesproducten in de etalage zetten: *tools* waarmee ook hun klant snel een ‘machine’ aan de praat krijgt om de goudklompjes uit de eigen databestanden te zeven.

Visser & Visser is een jong, middelgroot accountants- en advieskantoor dat vooral voor mkb-ondernemingen werkt. Partner Wilco Schellevis geeft leiding geeft aan een team van zeven mensen dat nieuwe, big data-gereleerde diensten ontwikkelt.

Vanuit een accountancy-achtergrond begon dat team software te bouwen en dashboards te bedenken waarmee die klant de informatie uit beschikbare bestanden kan ontsluiten. “We hebben hier in huis een soort standaardoplossing gemaakt die bij de klant, met wat kleine aanpassingen, meteen operationeel kan worden gemaakt. Daarmee heeft de klant toegang tot stuurinformatie zonder dat er vooraf grote investeringen moeten worden gedaan in mensen of softwareontwikkeling.”

Schellevis noemt een voorbeeld van een klant, actief in afvalverwerking, die snel rendement uit zo’n dashboard kon halen: “De directie kon ineens zien op welke plekken in het verwerkingsproces verspilling optrad, en hoe de zaak operationeel kon worden aangepast.”

### ‘Nog vrijwel niets’

Een vraag die opdoemt is of je hier wel echt van ‘big’ data sprake is. Want ‘exploratief’ zoeken in grote bestanden is niet aan de orde. Dat is precies de kanttekening die Pieter de Kok maakt. Hij is bekend als het ‘gezicht’ van de Tuacc accountantsbeweging, en in het dagelijks leven partner bij Coney, dat zich profileert als aanbieder van data-analyses, *process mining*-oplossingen en -diensten. De Kok: “Het potentieel van big data is enorm. Maar op dit moment gebeurt er nog vrijwel niets mee, zeker in de accountancy niet. Dat ligt aan datakwaliteit en -beschikbaarheid, investeringsbereidheid, en aan ontbrekende visie en business cases. Bovendien stuit je al snel op privacywetgeving.”

De Kok ziet wel ambitieuze toepassingen in ontwikkeling: “Ja, in de health sector bijvoorbeeld. Maar met big data problemen oplossen rond de goedkeuring van jaarrekeningen - zoals nu speelt in de ziekenhuizenwereld - dat gaat nog niet.”

De Kok ziet mkb-dashboards eigenlijk niet als big data-

## Big data analytics als wapen voor toezichthouders

Een interessante manier om big data in te zetten in een voor oob-accountants vertrouwde omgeving, de bestuurskamers, wordt dit voorjaar gelanceerd door oud-toezichthouder Paul Koster. Registeraccountant Koster was de afgelopen vijf jaar in Dubai toezichthouder op financiële instellingen gevestigd in de financiële freeport Dubai International Financial Centre (DIFC). Voordien was hij zeven jaar bestuurder bij de Nederlandse Autoriteit Financiële Markten.

Begin 2014 keerde Koster uit Dubai terug, en bekijkt nu de mogelijkheid van het starten van een consultancy die commissarissen helpt bij hun toezicht.

Koster: "Een zwak punt in het *two tier governance*-model van vennootschappen en andere *public interest entities*

vind ik dat de toezichthouders voor cruciale informatie te veel afhankelijk zijn van het uitvoerende bestuur. Ik denk dat die afhankelijkheid kan worden verminderd als raden van commissarissen zélf de nodige informatie gaan zoeken via big data analytics."

Door rechtstreeks bronnen aan te boren via data-generatie en data-analytics zouden de toezichthouders het management volgens hem betere vragen moeten kunnen stellen, bijvoorbeeld over marktontwikkelingen. Ook kunnen discussies met de directie over voorliggende plannen en de mogelijke consequenties daardoor meer diepgang krijgen.

De oud-toezichthouder ziet enorme onbenutte mogelijkheden, al beseft hij dat het er veelal op zal neerkomen dat big data moet worden terug geschaald tot *small big data*. Dus strakker gedefinieerd.

"Bijvoorbeeld toegespitst op een vergelijking van de concurrentie in een specifiek gebied. Of gerelateerd aan overname- of financieringsissues." Koster meent dat raden van commissarissen de directies zo beter kunnen assisteren en op relevante zaken kunnen anticiperen.



oplossing: "Er is immers niet echt sprake van 'volume'. Daarvan kun je misschien spreken als je een hele productieketen in beeld hebt, en daardoor nieuwe inzichten kunt bieden. Anders is het 'gewoon' data-analyse." Schellevis maakt zich over die definitie minder druk: "Moet het altijd gaan om nieuwe vergezichten en grote innovaties? Wat doet het ertoe? Wij geven mkb-bedrijven gewoon toegang tot een technologie die helpt om efficiënter te werken." De zelfontwikkelde dashboardoplossing Refine-IT wekt inmiddels ook interesse bij grotere bedrijven. Schellevis: "Daar zouden we als mkb-auditor anders niet zo snel binnenkomen." Een eerste teken van verschuivingen in het concurrentieveld?

### Impact op concurrentieverhoudingen

Een nieuwe technologie die het verdienmodel in een bedrijfstak overhoop kan gooien, heeft bijna altijd impact voor de concurrentieverhoudingen. En één vraag is daarbij bijna altijd cruciaal: is toegang tot kapitaal ja of nee doorslaggevend om van die technologie te kunnen profiteren?

De *move* van KPMG om een investeringsfonds op te zetten, geeft de indruk dat dit hier speelt. Anderzijds laat de aanpak van Visser & Visser zien dat investeringsdrempels relatief eenvoudig overbrugd kunnen worden en dat het voor mkb-accountants en hun klanten belangrijker is dat de wil aanwezig is om energie en aandacht in het onderwerp te steken. Dat komt ook naar voren in de opmerkingen van

## FOU-KHAN TSANG (ALFA): 'ALS JE DE DEFINITIE VAN 'BIG DATA' NIET AL TE STRAK FORMULEERT, DAN ZIJN ER ZEKER VOORBEELDEN VAN MKB-KANTOREN DIE DATA-BESTANDEN KOPPELEN EN DAAR INFORMATIE UIT HALEN.'

Fou-khan Tsang, voorzitter van Alfa Accountants en Adviseurs en bestuurslid van het netwerk van accountantskantoren SRA. "Als je de definitie van 'big data' niet al te strak formuleert, dan zijn er zeker voorbeelden van mkb-accountantskantoren die databestanden koppelen en daar informatie uit halen. Zo legde een SRA-kantoor dat de cijfers van een auto-importeur controleert, de ontvangen omzetgegevens naast data van de Rijksdienst voor het Wegverkeer en CBS-gegevens over geïmporteerde auto's. Dat bleek heel effectief."

### Benchmark-systeem

De SRA ontplooit vanuit haar backofficefunctie ook zelf initiatieven om te zorgen dat kleinere kantoren met 'big data' kunnen gaan werken.

Menno Kooreman, hoofd SBU automatisering bij SRA in Nieuwegein: "We bouwen nu een benchmark-systeem dat wordt gevoed met data van aangesloten

## Controllers: Rapporteren zonder grootboek

Het grootboek is de centrale bron van informatie voor controllers bij hun periodieke managementrapportage. Is dat nog logisch in een tijd waarin een bijna oneindige hoeveelheid data uit verschillende bronnen beschikbaar komt?

In het najaar van 2013 was in een VPRO Tegenlicht-documentaire te zien dat het mogelijk is om op basis van de bezetting van parkeerplaatsen bij retailketens de omzetontwikkeling redelijk nauwkeurig in kaart te brengen. Deze eenvoudige toepassing van data-analyse is op tal van plaatsen mogelijk. Neem de winkels op Schiphol: de omzet hiervan is vrij goed te voorspellen met data over het vluchtschema van de komende dagen. Want met ervaringscijfers weten we hoeveel meer een Rus gemiddeld uitgeeft dan een Nederlander.

Ongeveer. In dat woord zit de kern, aldus Sander Klous, bij KPMG verantwoordelijk voor een team dat adviseert over data analytics. Financials zijn volgens Klous opgevoed met het idee dat financiële informatie exact moet kloppen en zijn mede daarom 'verslaafd' aan het grootboek, waarin immers exacte waarden zijn opgenomen. "Een verkoopfactuur moet niet ongeveer goed zijn maar exact en vindt dus ook exact zijn weg naar het grootboek. Voor een ordelijke bedrijfsvoering is dat natuurlijk maar goed ook, want zo gaat de klant niet klagen en weet de debiteurenadministratie ook precies hoeveel geld er binnen moet komen. Maar de exactheid is niet erg relevant voor de directeur die wil weten hoe zijn bedrijf of business unit deze maand heeft gedraaid."

Een oude accountantswijsheid bij de waardering van zachte posten als immateriële activa is dat je beter ongeveer goed kunt zitten dan precies fout. Maar waarom wordt die wijsheid

niet toegepast als het gaat om zaken als kosten, omzet, marges, cash flow en dergelijke? Het is dan ook tijd dat controllers het juk van de exactheid afschudden, door meer te redeneren vanuit de informatiebehoefte bij het management dan vanuit het aanbod aan data vanuit het grootboek.

Controllers die deze omkering daadwerkelijk voor elkaar krijgen zullen waarschijnlijk (combinaties van) heel andere informatiebronnen gaan gebruiken. Variërend van marktdata en intelligente scans van social media tot en met betaaldaten en gps-patronen van de telefoons van hun medewerkers. Waarom ze dat zouden moeten doen? Omdat het sneller, gemakkelijker en beter is dan voortbouwen op het grootboek.

Zijn die data wel betrouwbaar genoeg om te komen tot een goede rapportage? Klous - tot enkele jaren geleden bezig om een niet te bevatten hoeveelheid data in de Zwitserse deeltjesversneller te doorzoeken op bewijzen voor het Higgs Boson deeltje - meent dat we ons daarover juist niet te veel zorgen moeten maken. "De hoeveelheid data groeit exponentieel. En hoe meer data er - uit verschillende bronnen - beschikbaar zijn, hoe beter de mogelijkheden worden om met wiskundige waarschijnlijkheidsanalyses vast te stellen of ze betrouwbaar zijn. Dat is ook precies hoe Google al jaren succesvol is, volgens het principe *Quantity over Quality*. De simpele redenering: alle fouten in data zijn op te lossen door er meer data aan toe

te voegen." Dat druist in tegen de intuïtie, maar Google bewijst toch steeds weer tot betrouwbare informatie (en voorspellingen) te kunnen komen met dat mantra. Het lijkt dan ook maar een kwestie van tijd voordat het mantra ook zal worden omarmd door vooruitstrevende financials die hun managementrapportages gaan bouwen op basis van een veelheid aan informatiebronnen.

In bepaalde sectoren helpt wetgeving daarbij een handje. Zo is de rode lijn in het nieuwe toezichtskader Solvency 2 voor verzekeraars dat zij (management)informatie moeten kunnen opleveren die een economische afspiegeling vormt van het bedrijfsbeleid en de keuzes die daarin worden gemaakt. De rapportage is meer toekomstgericht dan vroeger en bedrijfsrisico's worden expliciet gekwantificeerd in marktwaarden.

Dat alles wordt hoofdzakelijk bepaald door de karakteristieken van de populatie verzekerden. Data over de gemiddelde voedingspatronen, professionele achtergrond en afkomst van de populatie verzekerden geven dan ook aanknopingspunten voor de hoogte van de (toekomstige) uitkeringen. Sterker nog: zonder deze informatie tasten bestuurders in het duister of hun premies wel afdoende zijn voor een positief rendement.

En al die informatie is niet terug te vinden in het grootboek...

Nart Wielaard

### MENNO KOOREMAN (SRA): 'WE BOUWEN NU EEN BENCHMARK-SYSTEEM DAT WORDT GEVOED MET DATA VAN AANGESLOTEN KANTOREN.'

kantoren. Die voeren gecontroleerde jaarrekeningen in dat grote databestand in. Deze gegevens worden verwerkt en daarmee kunnen SRA-kantoren snel een beeld krijgen van wat in een bepaald type onderneming of een sector 'normaal' is."

Tsang: "Je kunt je voorstellen dat het voor een accountant die een controle doet in een notoir lastige branche

als de horeca, dankzij die benchmarkdata makkelijker is te zien wat opvallend is en wat niet."

In de visie van IT-specialist Kooreman draagt een parallelle technologische ontwikkeling er toe bij dat investeringen geen barrière hoeven zijn om met big data te werken: "Dat is het fenomeen van *the cloud*. Veel informatie en rekenkracht is nu tegen lage kosten via die weg te bereiken."

#### En toch zeurt er iets

De meeste toepassingen die onder het spannende kopje 'big data' worden geschaard, zijn vooralsnog dus behoorlijk aards. En Fouk Tsang erkent ook dat big data onder mkb-accountants nog niet een gevoel van strategische of praktische urgentie oproept. "Ja, er zijn een paar ambitieuze *early adopters*, maar ik zie eerder een

geleidelijk toenemen van toepassingen in de komende twee tot zeven jaar. Een meerderheid kijkt de kat nog uit de boom.”

En toch blijft er in het achterhoofd iets zeuren bij het aanhoren van de jubelverhalen over hoe big data *the world as we know it* gaat veranderen en verfraaien. Ja, privacy is natuurlijk een enorm probleem. Maar dat krijgt al volop aandacht. Zowel van degenen die een *big-brother-is-watching-you* maatschappij vrezen als van Google, Apple, Facebook en al die andere bedrijven en organisaties die begerig kijken naar de schat aan data over bedrijven en personen. Hoe meer privacywetgeving, hoe minder mogelijkheden om alle commerciële toepassingen te exploiteren (zie ook pagina 18).

Maar er sluimeren nog andere problemen. Bij het zoeken naar correlaties zal het regelmatig voorkomen dat bij onderzoek patronen oplichten die volgens de regels van de statistiek significant zijn maar vervolgens helemaal verkeerd worden geïnterpreteerd. Verkeerd in de zin dat aan correlaties ten onrechte een causaal verband wordt toegedicht.

#### 'Big errors'

De al eerder aangehaalde Nicholas Nassim Taleb, auteur van *The Black Swann* en *Anti-fragile* schreef daarover in augustus 2013 het stuk *The Big Errors of Big Data*. Boodschap: van de tien correlaties die je vindt zijn er misschien maar twee of drie de moeite van nader onderzoek waard. En hoeveel inspanning zal het niet kosten om uit te maken welke dat zijn?

Bovendien: het is niet ondenkbaar dat de programmeurs van big data *queries* in de valkuil storten waarin de zoveel macro-economische modellenbouwers als sinds 2007 liggen: ze hadden hun modellen zozeer verfijnd en er werd zoveel van hun berekeningen verwacht dat het idee om ook nog eens uit het raam te kijken naar de echte wereld niet meer opkwam. *Garbage in, garbage out!*

Een groot onderliggend probleem zit dus in het zorgvuldig definiëren van kenmerken van mensen of organisaties die in een databestand zijn opgeslagen. En natuurlijk van de 'algoritmen', de rekenkundige formules waarvan afhangt of een reeks kenmerken als een patroon wordt aangemerkt en verwerkt.

#### Intuïtie aspect

Sander Klous is hoofd van een nu nog vijftienkoppig team bij KPMG Advisory in Amstelveen genaamd 'Big Data Lab'. Dit team ontwikkelde de app die op 30 april - inhuldigingsdag van Koning Willem Alexander - via nu.nl aan het publiek beschikbaar werd gesteld om inzicht te geven in waar menigten zich in de stad concentreerden.

Klous erkent volmondig dat de opgeworpen mitsen en maren reële problemen zijn: "Opdrachtgevers van big data *queries*, *data analysts* en *data scientists* moeten zich daarvoor voortdurend bewust zijn."

Toch is hij het hartgrondig oneens met big data-sceptici zoals Taleb. "Het is goed dat hij waarschuwt, maar met

## Data-analyse en Process Mining

Ontdek de mogelijkheden van data analyse en process mining voor de (audit) praktijk en neem kennis van de laatste ontwikkelingen op het gebied van tools en technieken hierbij.

Datum: 15 mei 2014

Locatie: Amsterdam

Kijk voor meer informatie op [www.nbaopleidingen.nl](http://www.nbaopleidingen.nl).

OPLEIDINGEN

NBA

## SANDER KLOUS (KPMG): 'DE TRANSFORMATIE NAAR EEN DATA-DRIVEN MAATSCHAPPIJ IS IN VOLLE GANG. HET IS GEEN KEUZE MEER.'

zijn scherpe toon polariseert hij - en ik denk tegen beter weten in. Want de transformatie naar een data-driven maatschappij is in volle gang. Het is geen keuze meer." Klous weet maar al te goed dat er fouten kunnen worden gemaakt die soms te laat of slechts bij toeval worden ontdekt. "Het intuïtie-aspect kan bij deze technologie in de knel raken."

Ieder die uitkomsten van queries onder ogen krijgt, moet verdacht zijn op fenomenen zoals Simpsons paradox (verborgen invloeden kunnen analyseresultaten vertekenen) en andere meet- of interpretatiefouten.

Klous: "Zo verscheen onlangs het *Handboek vrouwspecifieke geneeskunde*, van ene professor Fauser.

Hij wijst erop dat bijna alle medicijnen die we nu gebruiken tientallen jaren geleden zijn getest, maar met mannen als proefpersonen. Die testresultaten zijn toegepast alsof ze ook voor vrouwen gelden. Ten onrechte, zo bleek inmiddels, en met schadelijke consequenties voor vrouwen."

De leider van het Big Data Lab meent echter dat dit type problemen kon worden opgelost door hooggekwalificeerde mensen met ieder hun eigen specialisatie bij elkaar te zetten, en een flexibele aanpak toe te passen waarmee ervaring is opgedaan in softwareontwikkeling: *Agile Development*.

Techneuten die definities en zoekcriteria moeten omzetten in digitale instructies, werken daarbij zeer nauw samen met de afnemers van de *queries*. Bij KPMG zijn dat advies- en auditteams, daarbuiten zouden het dokters, actuarissen of riskmanagers kunnen zijn.

Klous: "Met deze werkwijze proberen we die intuïtie weer in te bouwen. Vakinhoudelijke kennis en kennis van analyse vullen elkaar aan." Tegelijkertijd lijkt het hem goed de verwachtingen wat te temperen: "Hier staat nog geen machine waar de gouden eieren vanzelf uit rollen." □