



Big Data schept nieuwe mogelijkheden

Big Data heeft verschillende definities, afhankelijk van wie deze geeft. Marc Teerlink is global strategist bij IBM, hij geeft zijn definitie en gaat dieper in op alle mogelijkheden die ontstaan door Big Data.

"Big Data" betekent voor mij: veel. We vinden het heel gewoon dat opslagruimte voor data steeds groter wordt, en omdat we ook steeds meer op willen slaan valt dat niet echt op. Big Data betekent ook andere soorten data dan de gebruikelijke gestructureerde lijstjes, formulieren en databases. Denk hierbij aan data die een andere structuur heeft, afkomstig uit sensors, boeken, documenten, beeld, geluid, gps locaties, zelfmeting (hardlopen, bloeddruk). Enorme hoeveelheden data bewaren heeft weinig zin als je er niets mee doet. De waarde van Big Data zit hem juist in de analyses die er op losgelaten worden. Ik zie Data als nieuw productiemiddel, een delfstof die nooit op raakt waarbij het bedrijfsleven de informatie uit die data uitstekend kan gebruiken in de bedrijfsvoering. Om klanten beter te bedienen, om efficiënter te gaan werken en hierdoor meer winst te maken. Door het koppelen van data en slim filteren ontstaan tal van nieuwe mogelijkheden. Uitvinder van het World Wide Web Tim Berners-Lee, voorspelde jaren geleden al dat aan een digitaal raffinageproces ongelofelijk veel nieuwe dingen zullen ontspruiten."

Delen is het toverwoord

Belangrijk is echter of men in staat is al die data te ontsluiten zodat de informatie bruikbaar wordt. "Ontsluiten is de eerste stap en daar helpen filtermechanismen bij. Stel dat ik in een zoekmachine Paris Hilton intyp, dan krijg ik 46 pagina's met info over een dame met die naam. Maar ik wil informatie over een Hilton hotel in Parijs. Goede filtering kan voorkomen dat ik informatie voorgeschiedt krijg die ik niet zoek." Maar informatie moet ook zichtbaar gemaakt worden. "De smartphone is niet populair geworden door het leuke schermje, maar omdat het achterliggende programma's heeft

waar ik van alles mee kan doen: muziek downloaden en beluisteren, foto's maken en bewerken, snel informatie vinden."

Voorspellen is vooruit rijden

Oude systemen geven gemiddelde voorspellingen, gebaseerd op het verleden. "Geen goede zaak", meent Teerlink. "Uit onderzoek blijkt dat tweederde van alle bedrijven beslissingen neemt op basis van het verleden. Alsof je via de achteruitkijkspiegel ergens naar toe rijdt. Het is echter veel belangrijker vooruit te kijken naar waar je heen wilt, of je voldoende benzine hebt en hoe laat je ongeveer aan zult komen." Teerlink geeft daar een voorbeeld van. Een gemiddelde supermarkt met ook een aanbod aan kleine elektronische artikelen en een website met voorraadbeheer. "Op de site staat dat een artikel op voorraad is, maar eenmaal in de winkel blijkt het artikel er toch niet te zijn. De klant zoekt dan in de winkel via het internet naar een webshop en bestelt het dan daar. Deze supermarkt luisterde naar social sentiment over deze producten op het internet en combineerde dit met locaties. Dat zorgde ervoor dat de logistieke bevoorrading aangepast werd per specifieke locatie, waardoor het nog maar zelden voorkwam dat iets niet op voorraad was (of te veel op voorraad had). De bevoorradingsvoorspelling verbeterde met 24 procent, de winst steeg met 3 procent." Het gaat om het ontsluiten van verbanden en verbinding leggen met andere partijen om die verbanden winstgevend te maken.

Denken in mogelijkheden

Teerlink vindt Big Data zeker niet iets om bang van of voor te zijn. "Integendeel, ik zie volop kansen en mogelijkheden. Meer data betekent meer informatie en door die slim in te zetten, zijn er ontelbare mogelijkheden om efficiënter, beter en winstgevender te werken." Ook hier geeft hij enkele voorbeelden. "Een app waarmee men vanuit de eigen agenda een afspraak kan plannen bij de kapper, de tandarts, de garage. Een boekingsstelsel ontwikkelen tussen verschillende bedrijven,



Marc Teerlink
Global strategist bij IBM

"Meer data betekent meer informatie en door die slim in te zetten, zijn er ontelbare mogelijkheden om efficiënter, beter en winstgevender te werken"

het 'aan elkaar knopen' van informatiestromen." Big Data kan als ruwe grondstof bijdragen de economie omhoog te trekken, nieuwe diensten creëren en de digitalisering van de samenleving nieuwe vorm geven." Slim inzetten van data kan veel tijd en geld besparen. Teerlink: "Een groot onderzoek in de Verenigde Staten stelde dat de gemiddelde medisch specialist gemiddeld vijf uur per maand vakliteratuur leest. In een jaar loopt hij/zij 10.000 uur informatie achter, dat is niet in te halen. Die specialist heeft baat bij een informatiesysteem waar de benodigde relevante gecombineerde informatie voor zijn neus verschijnt terwijl hij met de patiënt bezig is. Er zijn talloze mogelijkheden om deze technologie te integreren."

Persoonlijke data vermarkten

Ook bij Big Data is privacy een belangrijk issue. "Het is verschrikkelijk belangrijk, maar mensen dienen zich te realiseren dat zij op allerlei manieren data genereren en delen. Of we nu spreken over het tweeten van de afstand en snelheid die je vanochtend gelopen hebt of het aantal filmpjes op Youtube dat je upload. Aan de ene kant is Big Data anoniem door zijn veelheid en de toepassing in het raffineren van patronen, trends en voorspellingen die zuiverder zijn dan een steekproef. Anderzijds moeten mensen bewust worden van die datastromen en beseffen dat het geld waard is. Het is heel goed mogelijk om de eigen data te vermarkten, het wachten is op bedrijven die hier handel in zien en dit voor jou tegen jouw prijs willen doen. Het belangrijkste is echter dat men ook NEE kan zeggen tegen gebruik van persoonlijke informatie. We kunnen deze ontwikkeling niet tegenhouden, het is dus verstandig alvast met zijn allen te bedenken hoe we er mee om moeten gaan in de komende jaren."

Nabije toekomst

Hoewel sommigen zich nog afvragen hoe alle informatie uit de toenemende stroom data toch ooit verwerkt kan worden, bestaat er sinds 2011 een supercomputer 'Watson' die een in spreektaal gestelde vraag kan interpreteren en na een zoektocht door 200 miljoen pagina's binnen enkele seconden een goed antwoord geeft op een vraag. De computer maakt gebruik van op maat gemaakte 'diepe-vraag-antwoordssoftware die veel verder gaat dan de conventionele kunstmatige intelligentie. "Deze computer is een kentering. Leerden wij hoe we met technologie om moesten gaan, nu leert de technologie hoe met ons om te gaan."

FEITEN & CIJFERS

BIG DATA

- ➔ **Datastromen veranderen** continu, het onderstaande overzicht is slechts een momentopname.
- ➔ **Zuckerberg stelt dat** er dagelijks één miljard stukjes content worden gedeeld via Facebook (status, vrienden, groepen, likes).
- ➔ **Eric Schmidt van Google** stelt dat we tegenwoordig elke twee dagen net zoveel nieuwe data maken als vanaf het begin van onze beschaving tot 2003.
- ➔ **Tachtig procent van** alle data is ongestructureerd: via sensors, weerpatronen, likes, commentaren, reviews en artikelen. Ongestructureerde data groeit vijftien keer sneller dan gestructureerde data.
- ➔ **Elke seconde wordt** er een uur aan video geüpload naar Youtube.
- ➔ **Wereldwijd worden** er elke 48 uur een miljard tweets verzonden. In maart 2013 was dit nog in drie dagen en in oktober 2012 een week.
- ➔ **Het record aantal tweets** was 456 per seconde, direct na het overlijden van Michael Jackson.
- ➔ **Zeventig procent van** alle data komt van individuele mensen, tachtig procent daarvan wordt door bedrijven opgeslagen en verwerkt.
- ➔ **Elke maand zijn** er meer dan honderd miljard zoekopdrachten op Google, in 2007 was dit nog 2,7 miljard.
- ➔ **Rond 2020 zal** een derde van alle data opgeslagen of door de cloud zijn gegaan en hebben we 35 zetabytes aan data gecreëerd.
- ➔ **Een bedrijf dat** de oorsprong van de big bag probeert te verklaren, verwerkt één exabyte aan signalen en sensordata per dag.
- ➔ **Meer dan tachtig procent** van de 247 miljard verzonden e-mails per dag zijn spam.
- ➔ **Verkeerde data** of data van slechte kwaliteit kost het Nederlandse bedrijfsleven jaarlijks 39 miljard. Het gaat dan voornamelijk om verkeerde adressen, verkeerde facturen, orders die geaccepteerd worden terwijl het artikel niet op voorraad is en een te optimistische kijk op openstaande bedragen.