



Heel groot en heel klein

Big data en quantum informatie

Aan zowel de groei van data als de groei van de verwerkingskracht van computers lijkt geen einde te komen. Maar groei alleen geeft nog geen goede besluitvormingsinformatie. Daarvoor is naast goede software, inzicht in gegevens en doelgerichtheid nodig.

DOOR: GUUS PIJPERS



In 1965 voorspelde Gordon Moore dat het aantal transistoren, en daarmee de verwerkingskracht van computers, elke twee jaar zou verdubbelen. Tot op heden geldt die wet nog steeds. Toch is het einde in zicht, omdat we met de huidige technieken omstreeks 2050 op het atomaire niveau zijn beland. Maar naast de tweejaarlijkse verdubbeling van Moore is de opslagcapaciteit zelfs exponentiël gegroeid. En hoe meer opslagcapaciteit, hoe meer gegevens je ook vindt om op te slaan (een groter huis vult zich ook bijna vanzelf met spullen). Voor zover mij bekend is er geen goeroe die een groeivoorspelling over data heeft gedaan. Verschillende onderzoeksbureaus hebben wel schattingen afgegeven van de groei van gegevens. In 2003 was er ongeveer vijf exabytes aan unieke informatie opgebouwd vanaf het begin van de mensheid. In 2009 was dat al opgelopen tot vijfhonderd exabytes, en in het najaar van 2010 produceerden we volgens Google elke twee dagen vijf exabytes informatie. Maar vergeet vooral niet dat nu al meer dan de helft van alle informatie een duplicaat is. Binnen vijf jaar is meer dan negentig procent van alle informatie een kopie. Hoeveel dat is? Vijfhonderd exabytes? Net zo veel als er zandkorrels zijn op de helft van alle stranden in de wereld. Of dertien keer een stapel boeken van de aarde tot Pluto. Als je al die boeken echt wilt drukken, heb je alle bossen van de aarde nodig. Twaalf keer maar liefst. En vijfhonderd exabytes is net zoveel wanneer iedereen op aarde een halve eeuw lang continu, 24 uur per dag, zou twitteren. Survival of the twittest!

■ Big data

De hoeveelheid gegevens groeit dus enorm. Die berg gegevens noemen we tegenwoordig big data. We spreken van big data als de omvang van de gegevens een onderdeel van het probleem is. Een grote hoeveelheid gegevens hoeft op zich geen probleem te zijn. Het KNMI of oliemaatschappijen bijvoorbeeld kunnen zeer goed met grote hoeveelheden data omgaan. Wat vaker aan de orde is, zijn goede vragen. Alle gegevens, klein of groot in aantal, wachten op een vraag. Een goede vraag stellen is namelijk vele malen moeilijker dan een antwoord geven. Zoals een anekdote laat zien: is

Aan groei van data lijkt geen einde te komen.

hoe vaak was Archimedes al in bad geweest voordat hij zijn Eureka-moment had bij de juiste vraag? En is de kroon van koning Hieron van Syracuse van puur goud?

Maar er is meer nodig dan een goede vraag. President Obama nam drie jaar geleden een belangrijke stap met zijn 'Open Government Initiative': alle overheidsinformatie online (data.gov). In een mum van tijd groeide het aantal beschikbare gegevens. Dit soort projecten kenmerkt zich altijd door enkele problemen. De originele gegevensbestanden worden niet geschoond of gestandaardiseerd alvorens ze online worden gezet. Verder worden gegevens vaak online gezet zonder een duidelijk doel of antwoorden op mogelijke vragen.

■ Quantum informatie

Onze wereld kunnen we in het groot en klein zien. Groot beschrijven we met de klassieke mechanica, de relativiteitstheorie. De kleine wereld van het gedrag van materie en energie op atomaire en subatomaire schaal beschrijven we met quantum mechanica. Een quantum systeem wordt beschreven met quantum informatie. Dat is de informatie die de verschillende toestanden van het systeem beschrijft. De belangrijkste informatie-eenheid is de quantum bit, meestal genoemd qubit. Een bit is de kleinste eenheid informatie die we kennen. Een bit heeft twee toestanden: 0 of 1. In de quantum mechanica rekenen we met quantum bits, qubits genoemd. Een qubit kan twee (of meer) toestanden tegelijkertijd aannemen. Elke toestand van een qubit is verbonden met een kans. Zo kun je bijvoorbeeld bij een qubit 25 procent kans hebben op toestand A en 75 procent kans op toestand B.



Daarbij is de totale kans van alle toestanden van het systeem natuurlijk 1. Bij een qubit weet je de werkelijke toestand (waarde) pas als je gaat meten. Stel dat je een dichte doos hebt met daarin een munt als voorbeeld van de genoemde qubit. Een onzuivere munt natuurlijk, maar dat doet er hier niet toe. Je rammelt met de doos. Op dat moment weet je niet welke kant boven ligt. Of anders gezegd, welke toestand de munt/qubit heeft. Wat je wel weet, is dat je bij opening van de doos in drie van de vier gevallen toestand B zult zien. Maar je hoeft de doos niet te openen, omdat je weet wat de toestanden zijn. Je kunt dus ook met meerdere dozen gaan rekenen zonder ze te openen!

We kennen allemaal de klassieke computer. Tegenwoordig in allerlei gedaanten, maar nog steeds volgens de principes van de Von Neumann-architectuur. Interessant is dat deze Von Neumann al meer dan een halve eeuw geleden werkte aan quantum mechanica, toen een relatief nieuwe discipline. Een computer die gebruikt maakt van de principes en uitgangspunten van de quantum informatie, wordt logischerwijs een quantum computer genoemd.

Quantum computers gaan in de toekomst een belangrijke rol vervullen. Een quantum computer van slechts 10 qubits kan 1024 berekeningen tegelijkertijd op één verwerkingseenheid uitvoeren. Dat noemen we *massive parallelism*, naast de bij klassieke computers bekende seriële en parallelle informatieverwerking. Deze quantum computers hebben echter een groot nadeel. Qubits moeten worden gesimuleerd met atomaire deeltjes, zoals een elektron of foton. Dat is in een laboratoriumomgeving nog wel te doen, maar nog lang niet in een smartphone of pacemaker. Maar net als bij de vergaande miniaturisering van de transistor geldt ook hier dat de juiste aandacht en inzet van technologie en experts tot slimme en innovatieve oplossingen leiden.

■ Encryptie

Quantum computers bestaan al sinds begin van de jaren tachtig, maar het eerste echte algoritme, factoring voor quantum computers, werd pas in 1994 geschreven. Factoring, ofwel het ontleden van een getal in zijn priemgetallen, is een van de

Quantum computers gaan een belangrijke rol vervullen.

belangrijkste praktische toepassingen. Factoring is de basis voor onder andere encryptie. Want hoe sneller dat ontleden gaat, hoe eerder je een bepaalde encryptie kunt ontcijferen. Elke encryptie die nu in gebruik is, gaat ervan uit dat dit kraken van de code met de huidige computers te lang duurt om effectief te zijn.

Een voorbeeld verduidelijkt dit. Stel dat je een getal tussen 1 en 16 wilt raden. Welke klassieke techniek je ook gebruikt, je hebt vier pogingen nodig om het juiste getal te raden. Anders gezegd voor een getal van 2^n heb je n pogingen nodig. Maar je kunt dit probleem via qubits oplossen. Dan heb je voor een getal van 2^n nog maar pogingen nodig. Dat scheelt voor grote getallen heel veel verwerkingstijd.

■ Business Analytics

Door een verregaande transparantie zien we dat miljoenen gegevens online beschikbaar zijn. Grote aantallen gegevens vragen om zeer krachtige computers. Het lijkt alsof de groei van de verwerkingskracht van computers de groei van data gemakkelijk kan bijhouden. Niets is minder waar. Want we hebben niet alleen computers nodig om gegevens te verwerken. We hebben die rekenkracht ook nodig om zinvolle, actiegerichte informatie te halen uit die grote bergen data. Naast natuurlijk softwarematige ondersteuning om onze, vaak latente of impliciete informatievraag beter te kunnen ondersteunen. Het nieuwste vakgebied heet dan ook Business Analytics, omschreven als het leveren van de juiste besluitvormingsinformatie aan de juiste persoon op de juiste tijd. Het vak is daarbij een vervolg op Executive Informatiesystemen en Business Intelligence. Met Business Analytics maakt ict hopelijk eindelijk de belofte waar om informatie sneller, persoonlijker en effectiever aan de besluitvormer te geven. ■

Literatuur

- Baeyer, H.C. von (2003). *Information: The New Language of Science*. Cambridge, MA: Harvard Publishing Press.
- Bokulich, A. & Jaeger, G. (2010). *Philosophy of Quantum Information and Entanglement*. Cambridge: Cambridge University Press.
- Clegg, B. (2006). *The God Effect – Quantum Entanglement, Science's Strangest Phenomenon*. New York: St. Martin's Press.
- Cox, B. & Forshaw, J. (2012). *The Quantum Universe: (And Why Anything That Can Happen, Does)*. New York: Da Capo Press.
- Floridi, L. (2010). *Information: A very short introduction*. Oxford: Oxford University Press.
- Jaeger, G. (2006). *Quantum Information: An Overview*. Berlin: Springer.
- Lloyd, S. (2006). *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*. London: Vintage Books.
- Nielsen, M.A. & Chuang, I.L. (2011). *Quantum Computation and Quantum Information: 10th Edition*. Cambridge: Cambridge University Press.
- Nixon, T. (2011). Quantum information. In: Ramage, M. & Chapman, D. (Eds.). *Perspectives on information*. New York: Routledge.
- O'Reilly Radar (2011). *Big Data Now – Current Perspectives*. Sebastopol, CA: O'Reilly Media.

OVER DE AUTEUR:

Guus Pijpers is managing director van Acuerdis en redacteur van TIEM. Hij geeft adviezen over de inzet en het gebruik van informatie voor individuen en organisaties. ask@guuspijpers.com

- Pijpers, A.G.M. (2012). 'Quantum informatie', *Informatie*, Jaargang 54, verschijnt binnenkort.
- Seife, C. (2007). *Decoding the Universe: How the New Science of Information Is Explaining Everything in the Cosmos, from Our Brains to Black Holes*. London: Penguin.
- Vedral, V. (2010). *Decoding reality: the universe as quantum information*. Oxford: Oxford University Press.
- Weinberger, D. (2011). *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York: Basic Books.
- Zikopoulos, P.C., Eaton, C., deRoos, D., Deutsch, T. & Lapis, G. (2012). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. New York: McGraw-Hill Osborne Media.

Advertentie

Nieuw bij Uitgeverij TIEM**DE KUNST VAN HET NIEUWE WERKEN**

Over Het Nieuwe Werken (HNW) is al veel geschreven. Maar hoe doen organisaties dat? Hoe implementeren ze nieuwe manieren van werken? En hoe ziet dat er dan uit? Gaat Het Nieuwe Werken alleen over huisvesting, ict, over hr-aspecten? Of is het meer?

De Kunst van het Nieuwe Werken laat zien dat HNW vooral een kans op organisatievernieuwing is. HNW stelt organisaties in staat om op nieuwe manieren verbindingen te maken met klanten en professionals. En dat is bepalend voor het behalen van succes.

De samenstellers van het boek, Paul van den Brink & Rosan Gompers laten in het boek 19 organisaties verhalen over hoe zij Het Nieuwe Werken zien en hoe ze daar mee omgaan. Leer van de implementatievalkuilen, lees de do's en dont's.

De Kunst van Het Nieuwe Werken is een staalkaart van de huidige werkpraktijk. Vandaar dat de volledige titel van het boek luidt: De Kunst van Het Nieuwe Werken|Ervaringen&dewerkpraktijk.

Te bestellen in de shop:

www.uitgeverijtiem.nl
Prijs € 32,50
ISBN 97890792006

