Publication: MAIL & GUARDIAN
Title: What's the big deal about big data?
AVE: 55789.28

Page: 32
Publish Date: 24 February 2017
Author: Tamsin Oxford

What's the big deal about big data?

Data science is working to find sense in all the information we can now gather

Tamsin Oxford

Big data is defined by research firm Gartner as high volume, high velocity and high variety information assets. It is generated by everything from social media and mobile devices to interconnected devices and machines. According to IBM, every day the world creates around 2.5 quintillion bytes of data and each byte is pulled from sensors, posts, images, purchases, GPS signals and more. It is a virtual swamp of information that can potentially be harnessed to transform lives, but it demands specialised analytics and technology, along with data scientists and experts, to ensure that the right insights are gleaned in meaningful ways.

"There is a distinct difference between traditional statistics and big data. Traditional statistics can be defined as a dam of known size and limited dimensions, and limited data loss or gain," says Tim Shier, founder and chief executive of BurningBennu. "Big data, on the other hand, is more like a river full of fresh data. The data is valuable for a short amount of time and you have no idea what is coming down the river next."

Where big data is defined as the data sets that are too large and complex for traditional interpretation to produce an outcome, data science is the field in which scientific methodologies are used to extract insight from big data. As technology evolves, so does the amount of data generated; it is increasing at the same pace, pouring forth from devices and systems without structure. The challenge is to find ways of storing and utilising the data to squeeze as much value from it as possible before it loses relevance.

"Data science has been rated as the sexiest career of the 21st century by the Harvard Business Review," says Katherine du Plessis, data analyst, Business Intelligence Unity, Magna Carta. "It is the fastest growing industry in technology because data is growing faster than we can manage it."

How do you solve a problem like big data? The data footprint of the individual is skyrocketing — some shared consciously over social media, some unconsciously through sensors, GPS signals and other data-heavy applications and solutions. Those with Google accounts can instantly see a snapshot of their timelines and travels by visiting *https://www.google.com/maps/timeline*. Still with Google, the human race enters around 40 000 new searches every second — a staggering total of 1.2 trillion searches per year, and rising. According to the International Data Corporation (IDC), 1.7 megabytes of new information will be created for every person on the planet, every second of every day, by the year 2020. The company also predicted that by then the data created will reach 44 trillion gigabytes — as much data as there are stars in the universe.

"People do have a growing digital footprint, and it raises many issues around data management and governance," says Jon Tullet, research manager, IT Services, IDC South Africa. "Organisations are already grappling with the volume and speed of their data while strategically investing in technologies to increase that data flow by orders of magnitude."

Analysis

The numbers for the data are so large that they almost defy conversation or comprehension. How can technology catch up with the river effectively enough to offer up real insight? There remains a clear challenge to find ways of analysing the data which are current, relevant and accurate. And there are some spectacular examples of organisations meeting this challenge head on.

The British Olympic rowing team has been using analytics to help their teams go faster and, according to *Forbes* magazine, currently they are the only British team to have won gold in every Olympics since 1984.

IBM has been using insights created by IBM Watson, a cognitive computing system that uses big data to help organisations make big decisions. Their list of successes is impressive, but is anyone else achieving anything meaningful from all the data?

"The answer is — yes, definitely," says Paul Stemmert, senior consultant, Delv Data. "The more data you have on a user's behaviour, interests etcetera, the more you're able to customise their experiences across digital channels. What's important to note here is that nobody ever got anything meaningful from analysing for analysis' sake. You only get meaning if you ask something meaningful of the data."

Dr. Yudhvir Seetharam, head of analytics at FNB, is not convinced: "Maybe. The type of industry tends to determine the maturity of the analytics. Some industries are more advanced, like financial services, while others are lagging behind, like mining or construction."

Before value can be eked out of big data, you have to know the questions to ask, and have the right tools and skills to achieve it. Meltwater, a global media intelligence provider, takes data from billions of conversations, articles, pages and mentions, and turns this into filtered insight so companies can act on it. They follow the trails of digital breadcrumbs to pull out specific and detailed information which the business can use to make decisions, shift strategy or engage with customers.

"Big data can be used in real time to tackle pressing business challenges and make smarter strategic decisions, but it is only useful if its insights are available to the right people at the right time," says Claude Schuck, regional manager, Africa, Veeam.

In a recent report, research firm Forrester estimated that the global spend on business intelligence and analytics software will grow at an annual average of 12% a year, reaching a global average of $33-billion by 2020. These are impressive numbers, especially if one is considering developing a solution to allow for the extraction and analysis of big data.

Big science?

Of course, big data doesn't just flow from the social, the individual and the internet of things; there are some remarkable scientific installations that are generating vast quantities of data such as the upcoming SKA and the Large Hadron Collider. The question is: how can this data be harnessed to gain insights that benefit humanity? This is where the rather dashing realm of data science kicks in, as South Africa seeks data scientists who can work on the SKA data to extract meaningful results. Several South African universities are highlighting the significance of big data and the need for data scientists, and this is being driven by the need to recruit hundreds of data scientists to meet the manpower demands of the SKA.

"The SKA and Large Hadron Collider are examples of why big data is needed to answer big science questions and why data science is a rapidly growing field," concludes Riaan Sunny, data analyst, Mark1 Media. "These projects require the storing, transferring and analysing of almost innumerable amounts of data that require the processing power of millions of computers and the diverse skill sets of data scientists. We are now living in the beginning of an era that could be likened to the start of the Terminator universe or even The Matrix, where computers are teaching themselves to become smarter, and we have big data to thank for it."

Claude Schuck, regional manager for Africa at Veeam. Photo: Supplied

This page was sponsored by the department of science and technology. Contents were sourced independently by the M&G's supplements editorial team.