

Welcome to the yotta world

Ludwig Siegele

Big Data will flood the planet

Even if you still have to think twice about the meaning of “giga” and “tera” in computer-speak, you’d better get ready for “peta”, “exa” and “zetta”. These binary prefixes, which denote 1m, 1 billion and 1 trillion gigabytes respectively, will be used more and more often: the amount of digital data is exploding, with a staggering 1.8 zettabytes in 2011, up from 1.2 zettabytes in the previous year.

This deluge has long been a topic among geeks; 2012 will be the year in which the Big Data trend gets noticed beyond their ranks. Many more firms will start to analyse huge piles of data to optimise everything from their supply chains to their customer relationships.

In the past, because data storage was expensive, lots had to be thrown away. Information was locked up in computer systems and could not be combined with other sources. Even if firms were able to aggregate their data in one place (called a “data warehouse”) and sift through the information, it was often already out of date.

All this is changing—rapidly. The price for storage is plunging; by 2020 storing a petabyte will cost a mere \$4, predicts Forrester, a market-research firm. Software to handle such huge amounts of data is improving, too. Hadoop, a new type of database, can be used to sift through big data streams in real time—and not just orderly numbers, but also “unstruc-

tured” data, for instance any kind of text.

In a world oozing data, sensors measure everything from speed to smell. Smartphone applications generate vast quantities of “data exhaust”: information that is produced when users engage in, say, skiing, and, thanks to a motion sen-

Talent to analyse data is scarce and will become scarcer

sor, the phone can tell whether somebody has just had a bad fall.

Social media are adding to the flood: the number of messages on Twitter will exceed 500m per day by the end of 2012. By opening their data vaults, some national and local governments are doing the same.

Most important, firms will discover

that they can extract value from the data. A study by the McKinsey Global Institute, a think-tank, found that analysing health-care data could yield \$300 billion-worth of savings in America alone. One of the pioneers in this field is Britain’s National Institute for Health and Clinical Excellence, which uses large datasets to investigate the cost and benefit of new drugs and existing expensive treatments.

Firms will also use data for new business models. Rolls-Royce no longer needs to sell all its jet engines; it can charge for their use. By continuously assessing their performance, it can predict when engines are more likely to fail, so that customers can schedule engine changes.

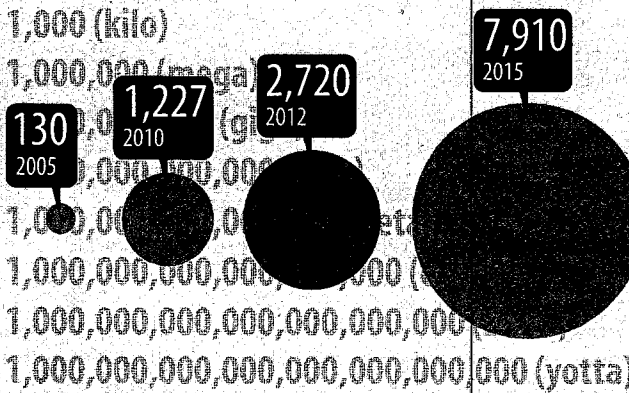
Already, the number of data-driven start-ups is growing—particularly when it comes to social media. Lexalytics, for instance, analyses the sentiment of utterings on Twitter, Facebook and other such services. Klout measures the influence of social-media users (and some firms give people with a high Klout score preferential treatment).

For Big Data to become huge, however, there are still hurdles to leap. For one thing, the tools to analyse data are not yet good enough. And people with the skills to analyse data are scarce and will become scarcer. By 2018 there will be a “talent gap” of between 140,000 and 190,000 people, says the McKinsey Global Institute. The main problem may be privacy. It is unlikely that people will want to live in a “yotta world” (1,000 trillion gigabytes) in which their every move is instantly digitised and added to the flood of public data. ■

Ludwig Siegele: online business and finance editor, *The Economist*

Exponential

Quantity of global digital data, exabytes



Source: EMC/IDC Digital Universe Study, 2011

2012 IN BRIEF

Gap, a clothing retailer, opens 30 new stores in China, while implementing the closure over two years of a fifth of its American stores



to share their locations, with a view to commercialising this information. But location information is most useful when tied to other things—such as the photos, status updates and recommendations shared on social networks. Plenty of action is likely in 2012 as the location start-ups both compete and co-operate with the big social networks (Facebook, Twitter and Google+), which will continue to integrate location information into their services.

A similar situation pertains in the field of augmented reality (AR), the sci-fi trick of overlaying information from the internet onto a real-time view of the real world. At the moment this is a party trick that can be performed with a handful of smartphone apps (such as Layar, Wikitude and Google Goggles) and on

some games consoles (such as the Nintendo 3DS). But as location and social-network information are knitted together, AR becomes the logical way to display the results: highlighting your friends in a crowd at a festival, for example, and eventually enabling your glasses to remind you of the name of the person you are talking to at a conference, together with their most recent postings on social networks. Once again this is enticing new territory that both start-ups and the big social networks have in their sights, along with search engines and hardware vendors. At the moment Apple is the company to beat when it comes to packaging cutting-edge technology in an elegant, easy-to-use form. Will Apple some day launch a set of augmented-reality goggles—iGlasses, perhaps? ■