

# Information Management

**Martin White, Information Focus**

## **Big Data – an information management perspective**

Big Data is a term now being widely used to describe the evolution of datasets that are beyond the ability of current database software tools to capture, store, manage, analyse and synthesis. If you read the countless pronouncements by the IT industry and by many consultancies with a vested interest you would be forgiven from thinking that unless your organisation invests in Big Data technologies there is no hope for the future. This column is a bit of a reality check!

### **What is Big Data?**

There are four characteristics of Big Data, volume, velocity, variety and value, a framework originally developed by the Gartner Group in 2001. There is nothing new in this world!

Volume - not only is the rate of addition of new information rapidly increasing through (for example) sensors in a piece of machinery, elsewhere the low cost of storage and lack of plans for archiving means that information is never discarded. One of the seminal reports on Big Data

was published by the McKinsey Global Institute in 2011 (see Resources) and the research carried out for this report in conjunction with International Data Corporation indicated that companies with 1,000 or more employees had at least 200 terabytes of stored data, whilst in many sectors (such as manufacturing and financial services) the volume of data was nearer one petabyte.

Velocity - the processes around updating databases of this scale in anywhere near real-time present substantial challenges in addition to the need to analyse this data in real-time given the very rapid changes in markets as a result of instabilities in the Eurozone and elsewhere. Even running standard reports across data that changes with such rapidity could lead to decisions being taken on the basis of outdated information, with consequential impacts on business performance and corporate reputation.

Variety - the need to interrogate information that is stored in multiple applications, perhaps using different versions of the same software and integrate this with, for example, quarterly re-

ports stored in SharePoint for presentation to a senior management team in PowerPoint is a substantial challenge. However the presentation may actually have been an hour-long videoconference and tracking down the video file adds a new dimension to the range of file formats being handled. Email traffic and social media applications will also contain a substantial amount of commentary and insight into operational data.

Value - very few organisations are able to quantify the volume, velocity and variety of the data they process, and none have a measure of its value, though for certain there are some benefits from being able to manage data more effectively.

## The business case for Big Data

Some of the categories of data that organisations are starting to see as offering potential for analysis with Big Data applications include:

- Financial transactions
- Logistics and traffic monitoring operations
- Quality monitoring from sensors on equipment and networks
- Retail transactions
- Social media analysis
- Telecommunications records
- Citizen data from government 'open data' sources
- Exploration data from natural resources exploitation

## Data and information

From the outset it is important to appreciate that Big Data is not just about database management though you would not think so from the hype which surrounds this topic. Collections of data only make sense when placed in the context of information about a customer or a market. Being able to analyse data on poor product reliability is of no value unless the circumstances around the product are known, such as the supplier, the way the product is installed and the service history.

Part of the folk law of corporate information is that 80% is unstructured and 20% is structured. Until earlier this year there was no research evidence for this statement but in May 2012 MarkLogic published a report that assessed the value of unstructured information and AIIM also released a survey that looked at the requirement to integrate unstructured and structured information. In the AIIM survey 60% of respondents said it was business-critical to be able to do so, but only 2% said that they could do so. There is so much useful information in these two reports that I would recommend you download them and read them in detail.

## The opportunity for information professionals

Senior business and IT managers are being overwhelmed by the current level of information, hype and misinformation about the importance of being able to manage Big Data. As information professionals we have a very significant role to play in monitoring developments in Big Data, no

matter what the size of the organisation, and in particular ensuring that our organisations develop information management strategies that enable information and data to be integrated. It is not just about the technology but about quality. On their own neither tell the full story, and of course even the best collection of information needs to be updated with information from external sources. Big Data will not be the end of the line for information professionals, but the start of a new era.

## Resources

### Reports:

- Big Data - the next frontier for innovation, competition and productivity

McKinsey Global Institute, 2011

[http://www.mckinsey.com/insights/mgi/research/technology\\_and\\_innovation/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation)

- Demystifying Big Data

TechAmerica Foundation, 2012

<http://www.techamericafoundation.org/bigdata>

- Big Data - Harnessing a Game-Changing Asset

Economist Intelligence Unit, 2011

<http://www.managementthinking.eiu.com/big-data.html>

### Market Surveys:

- Big Data - extracting value from your digital landfill

AIIM (<http://www.aiim.org/Research-and-Publications/Research/Industry-Watch/Big-Data-2012>)

- From Overload to Impact: An Industry Scorecard on Big Data Business Challenges

Oracle, 2012

<http://www.oracle.com/us/industries/industry-scorecard-1683398.html>

- Big Data is real and it is here

MarkLogic, 2012 <http://info.marklogic.com/post-relational-reality-dbta-survey-2012.html>

*Martin White is Chair of UKeiG and Managing Director of Intranet Focus Ltd*