Deloitte.

# Billions and billions: big data becomes a big deal

Deloitte predicts that in 2012, "big data"[27] will likely experience accelerating growth and market penetration. As recently as 2009 there were only a handful of big data (BD) projects and total industry revenues were under $100 million. By the end of 2012 more than 90 percent of the Fortune500 will likely have at least some BD initiatives under way. Industry revenues will likely be in the range of $1-1.5 billion. But the industry is still in its infancy. Big data in 2012 will likely be dominated by pilot projects; there will probably be fewer than 50 full-scale big data projects (10 petabytes and above) worldwide.

Historically most of the world's accessible data has been located in traditional relational databases, accessed and managed with a certain set of tools and analyzed and reported on with business intelligence software. The ability of those tools and software applications to cope with larger and larger data sets has grown over time, but in general any set of data that was viewed as being "too big" or needing results "too fast" was seen as requiring an entirely new set of tools, most commonly referred to as big data tools.

Until the last year or two, traditional data tool capacity had more or less managed to keep pace with the growth in data sets. However social networks, real time consumer behavior, mobility, sensor networks and other data generating sources have caused many organizations' data warehouses to overflow. Data sets an order of magnitude (or two) larger than before are either happening today, or are seen as likely within the next 12 months. Even when the size of a data set has not grown so quickly, if organizations now want to do analysis in real-time, sometimes traditional tools are not adequate and big data is again being considered.

From being the sort of tool that was only needed for meteorology or physics simulations, big data has recently moved into the mainstream: individual big data conferences[28] are drawing thousands, BD companies are attracting funding rounds of over $50 million[29], BD venture funds are being created[30], and large existing software players are validating the markets by partnering with or acquiring outright early stage leaders in the space[31].

Not all industries are likely to benefit from big data projects equally, and uneven distribution of BD across verticals is already perceptible. Not surprisingly, the first movers were Internet companies: in fact, the most popular big data tools are being built on top of software that was originally used to batch process data for search analysis[32]. The fast follower sectors are likely to be public sector, financial services, retail, and entertainment and media[33].

Like many emerging technologies, BD can be easy to read about, but hard to picture in real world applications. A recent example from the media industry may help:

*"The Financial Times uses big data analytics to optimize pricing on ads by section, audience, targeting parameters, geography, and time of day. Our friends at the FT sell more inventory because the team knows what they have, where it is and how it should be priced to capture the opportunity at hand. To boot, analytics reveal previously undersold areas of the publication, enabling premium pricing and resulting in found margin falling straight to the bottom line.[34]"*

27 Big data is a term applied to data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time. Big data sizes are a constantly moving target currently ranging from a few dozen terabytes to many petabytes of data in a single data set.

28 Enterprise Applications: Hadoop World 2011: 15 Key Takeaways, Slide 2, eWeek, 15 November 2011: http://www.eweek.com/c/a/Enterprise-Applications/Hadoop-World-2011-15-Key-Takeaways-566207/

29 'Big Data' Firm Raises $84 Million, The Wall Street Journal, 14 September 2011: http://online.wsj.com/article/SB10001424053111903532804576569133957145822.html

30 Accel Makes Big Commitment To Big Data With $100M Fund, The Wall Street Journal, 8 November 2011: http://blogs.wsj.com/venturecapital/2011/11/08/accel-makes-big-commitment-to-big-data-with-100m-fund/

31 Teradata to introduce new 'big data' analytic solutions to Egypt, Daily News Egypt, Dec 13, 2011 http://thedailynewsegypt.com/it-a-telecom/teradata-to-introduce-new-big-data-analytic-solutions-to-egypt.html

32 Hadoop: The Definitive Guide, Tom White, 2009, O'Reilly Media, Inc., http://db.trimtabs.com:2080/mindterm/ebooks/Hadoop_The_Definitive_Guide_Cr.pdf

33 This assessment is based on publicly disclosed big data project announcements and Deloitte member firm interviews.

34 How Big Data Analytics Can Save Publishing, AdAgeDigital, 5 December 2011: http://adage.com/article/digitalnext/big-data-analytics-save-publishing/231363/

35 Sizing the big data problem: 'big data' is the problem, The 451 Group, 9 December 2010: http://blogs.the451group.com/information_management/2010/12/09/the-big-data-problem/

36 Global IT Spending 2011 Projections Revised Upward to $3.7 Trillion, IT channelPlanet, 11 July 2011: http://www.itchannelplanet.com/trends/article.php/3937061/Global-IT-Spending-2011-Projections-Revised-Upward-to-37-Trillion.htm

37 Big Data consolidation march continues: Teradata buys Aster Data, OnStrategies Perspectives, 3 March 2011: http://www.onstrategies.com/blog/2011/03/03/big-data-consolidation-march-contines-teradata-buys-aster-data/

Like many nascent IT spaces, big data has been difficult to size[35]. Estimating the market size is challenging for several reasons: there are varied definitions of what BD is, it is still early in the adoption cycle of big data technologies, and most of the companies who are doing BD do not disclose their spending. Another barrier is that BD work is primarily based on open source code: the initial software is free, and the real spending comes from internal IT staff adapting the code. Unlike measuring sales of a new kind of router, BD spending is not easy to count.

Assuming moderate growth in 2012 over 2011, global spending on all information technology is projected to be roughly $3.7 trillion[36]. Of that, total enterprise software should be about $270 billion. And of that, database management systems spending (DB) will likely be over $27 billion, enterprise resource planning (ERP) software about $25 billion, and business intelligence (BI) roughly $17 billion. Combining the DB, ERP, and BI markets the total could be about $70 billion.

According to some sources, more than 90 percent of analytic systems/data warehouses have less than five terabytes of data, and can be handled by traditional database tools and analytics[37]. This would suggest that the value of the DB+ERP+BI requiring big data solutions would be at most $7 billion in 2012. Further assuming that it is still relatively early in the adoption cycle, 15-20 percent of that total might be converted in 2012, which drives a range of roughly $1-1.5 billion[38].

**Bottom line**

Traditional DB, ERP and BI companies will likely not be hurt in 2012, or even the longer term: big data will coexist with the existing DB management stack. At a recent BD conference, a survey showed that over 60 percent of respondents believed that their "Existing data warehousing/BI analytics supplier…will deliver big data technologies and solutions.[39]"

Further, just having the BD tools isn't enough, enterprises need to know what questions to ask, actually ask them, and then translate that into strategy or tactics. Moreover, a recent survey of chief marketing executives found that "…more than 60% of knowledge workers at large enterprises say their organizations lack the processes and the skills to use information effectively for decision making[40]."

Even though BD is still in its early stages, the growth suggests that the industry needs to develop talent with big data skill sets: 140,000 to 190,000 skilled BD professionals will be needed in the US alone, over the next five years[41].

It will be important for enterprises to develop new policies around privacy, security, intellectual property, and liability. BD isn't just about technology and employees with the right skill sets, it will also require businesses to align work flows, processes and incentives to get the most out of it[42]. It is important to note that enterprises should not concentrate on big data at the expense of "current data", or business information as normal. There is still a lot of value left to be extracted from the information inside their traditional databases!

There are plenty of reasons to be skeptical about the BD market. But Big data probably deserves a place in overall enterprise IT strategy for generating business insight. Best practices include generating a list of important challenges or questions that the current approach to data does not address. Could big data deliver the answers enterprises are looking for? If so, then it's all about discipline. A disciplined, targeted approach to big data – one focused on answering very specific questions for the business – is one that many companies can probably take on today without abandoning their current efforts[43].

Big data's potential is likely to pivot on context: when organisations recognise that big data's ultimate value lies in generating higher quality insights that enable better decision making, interest and revenues should accelerate sharply.

38 As a fact check, we have interviewed a sample of enterprises that are budgeting for 2012 Big Data projects. Average spend per company is about $10-15 million. Assuming that about 1000 companies do similar sized pilots also yields a $1-1.5B market.

39 Hadoop World 2011: The State of Big Data Adoption in the Enterprise, Slide 20, Slideshare, November 2011: http://www.slideshare.net/cloudera/the-state-of-big-data-adoption-in-the-enterprise-tony-baer-ovum

40 Beware The Big Data Hype, Forbes, 11 September 2011: http://www.forbes.com/sites/patrickspenner/2011/11/09/beware-the-big-data-hype/

41 Big data: The next frontier for innovation, competition, and productivity, McKinsey, May 2011: http://www.mckinsey.com/mgi/publications/big_data/

42 Big data: The next frontier for innovation, competition, and productivity, McKinsey, May 2011: http://www.mckinsey.com/mgi/publications/big_data/

43 Is Big Data Just a Big Distraction?, Deloitte: http://www.deloitte.com/view/en_US/us/Insights/Browse-by-Content-Type/deloitte-debates/9f372efd73310VgnVCM3000001c56f00aRCRD.htm



Designed and produced by The Creative Studio at Deloitte, London. 16470A