THE INTERNATIONAL JOURNAL OF BUSINESS & MANAGEMENT

Big Data: It's Actually Big

Rinku Mahindru
Assistant Professor, Department of Commerce
Faculty of Commerce and Business, Delhi School of Economics, University of Delhi, India

Abstract:
The term, Big Data first emerged on world stage in early 2000s when companies started to find it difficult to manage information data sets whose size grew beyond the ability of software tools managing data during that period. Today, in this ever competitive environment, Big Data has become a business priority for companies in the globally integrated economy. Big Data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making. In every Industry and business, data has become as vital as factors of production, i.e. labour and capital. A large pool of data is available globally, however if it is not accurately sourced, cleaned, integrated, or made available through system access, then it restricts management's ability to make timely, information-driven business decisions and achieve competitive advantage. Big Data not only provides solutions to long-standing business challenges, but also helps in finding new ways to transform processes, organizations, entire industries and even society itself. Big Data analysis drives nearly every aspect of the modern society, including mobile services, retail, manufacturing, financial services, life sciences, and physical sciences. This article emphasize on the meaning, techniques & technologies, trends and challenges of Big Data. It relates the concept of Big Data to DIKW model. It also details out the current scenario and future outlook of Big Data in India and the world at large.

Key words: Data, Big Data, information-driven, DIKW Model, competitive advantage

1. Introduction

The Big Data is a term used to express the voluminous amount of unstructured and semi-structured data in an organization that would take too much time and money to load into a relational database for analysis. Although the Big Data doesn't refer to any specific amount, the term is often used when speaking about petabytes and exabytes of data.

- Defining Big Data via the Three Vs: (TWDI, 2011) Most definitions of Big Data focus on the size of data in storage. Size matters, but there are other important attributes of Big Data, namely data variety and data velocity. The three Vs of Big Data (volume, variety, and velocity) constitute a comprehensive definition, and they bust the myth that Big Data is only about data volume. In addition, each of the three Vs has its own ramifications for analytics
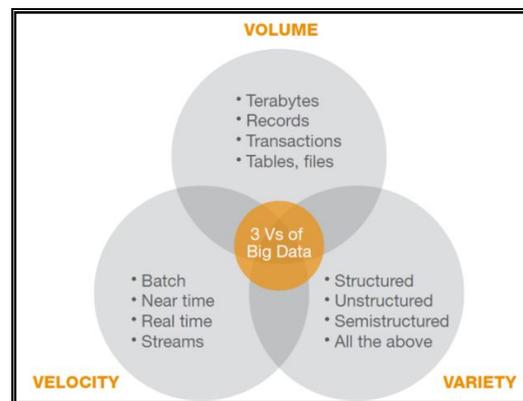

Figure 1: Defining Big Data via Three V's
Source: TWDI Research report on Big Data Analytics

2. Objectives
- To define the concept of Big data.
- To Identify the need for Big Data Analytics
- To highlight the challenges for Big Data specifically for business leaders, business organizations and IT organizations.
- To study the Big Data trends-2013
- To create value of Big Data by relating it to DIKW pyramid
- To identify the impact of Big data on various industries namely; Education, Healthcare, financial services, banking, manufacturing, insurance, communication etc.
- To throw light on Indian Big Data Scenario.
- To highlight the Current Global Scenario & Future outlook of Big Data.

3. Research Methodology
The research is based on the information collected from various secondary sources.
- Books, journals, working and white papers are used for reviewing the academic literature on the topic.
- Newspapers, on line news channels, industry reports and websites are accessed for the collection of any other required information related to the topic.

4. Need for Big Data Analysis
In this era of Globalization, vast amount of information is available that were neither available nor accessible in past. Earlier surveys, focus groups, etc. were some of the sources of information; however now they can be examined directly. Organizations can discover more through larger samples and crude information. This potential can be realized only if you bring together and analyze all that data. At present, we have gigantic information for analysts to manage, which increases the chances of missing opportunities or risks. Organizations that augment their human experts with Big Data technologies could have competitive advantages by heading off problems sooner, identifying opportunities earlier, and performing a mass customization at a larger scale.

5. Challenges of Big Data
Like all major technological innovations, Big Data has its own bricks and motors. Following are the challenges faced by the business leaders, business organizations and the IT organizations in adopting Big Data.

5.1. Big Data A Challenge For Business Leaders
- Businesses, consumers and suppliers are creating and consuming gigantic amounts of information in the global marketplace.
- Gartner predicts that enterprise data in all forms will grow 650 percent over the next five years.
- According to IDC, the world's volume of data doubles every 18 months. This flood of data, often referred to as "information overload," "data deluge" and "Big Data," clearly creates a challenge for business leaders.

5.2. Big Data A Challenge For Business Organizations
- Most organizations are either getting on a technique related to Big Data or intend to do so in the near future, yet almost no organizations have an expressed strategy for it.
- Big Data initiatives widen an organization's IT setup in new ways and can strain dealings with business units.
- Big Data initiatives originate within business units, putting added pressure on organization's IT department to get adequately prepared to support them.

5.3. Big Data A Challenge For IT Organizations
- IT organizations should understand that Big Data is not only big volumes. Big Data projects might not succeed if attention is not paid to variety, velocity, complexity, as well as volume.
- IT departments should be well prepared for things like budget changes and infrastructure re-engineering, as priorities will shift to Big Data processing and analysis.
- IT organizations should be on view that apart from some successful Big Data implementations in well known organizations, most Big Data projects in recent past are best-described as exploratory.
- IT leaders must make sure that appropriate budget and training should be provided as Big Data requires experienced and skilled IT personnel's.

6. Big Data Trends ("IT Market Trends and Predictions on Big Data", 2013)
Big Data is no longer a niche, having matured over the years into mainstream. It is everywhere – sports, healthcare, retail and the security industry to name a few areas. Some of the key trends include:
- **M&A**: Big Data technology and services will reach a benchmark primarily driven by M&A activities in predictive analytics, rich media analytics, and data discovery.
- **In-memory/real-time enterprise**: In-memory computing will become a key lever for organizations to differentiate and gain competitive advantage over the next 3 years.

- **Industry-specific offerings**: With the evolution of different use cases around leveraging Big Data," there is a renewed focus on industry-specific offerings.
- **Volume**: Large volume also implies more people to model and mine the data; data scientists will emerge as a critical role in enabling effective analytical outcomes.
- **Hadoop**: Increased data-mining demands on large sets of data coupled with the need for cost effective solutions have driven demand for the Hadoop/ Map Reduce software ecosystem.
- **Governance**: Big Data is becoming more of a mix of structured and unstructured data. With the proliferation of heterogeneous data, it is imperative to have transparent data governance standards, including data quality benchmarks, security rules, and clear stewardship.

7. Creating value through Big Data-DIKW Pyramid

When it comes to creating value we relate Big Data to the DIKW pyramid which is the progression from data to wisdom. Data alone is a mere building block. It can generate information based on facts. Interpretations drawn from it form knowledge, which is basically understanding of what the information implies. Finally, the knowledge is evaluated against experience, and if it's found to be useful it becomes wisdom.
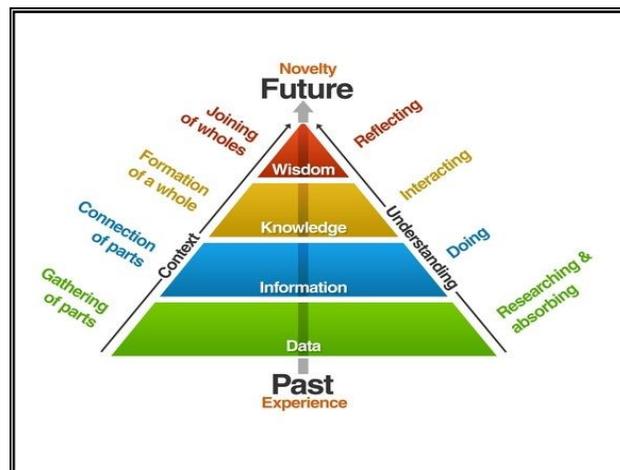

Figure 2: Big Data and DIKW
Source: KarimVaes(2013)

According to Gayle, Richard (Spreading Science),

- **Data** by itself has no meaning. It fills databases. This data must be examined in order to become meaningful. Generating more data accomplishes little. Humans must interact with it. This is the first step whereby humans and their social networks become important.
- **Information** arises when humans examine the data. This provides a framework for understanding what the data represent. This information can be tacit, held inside our heads, or explicit, presented in a fashion that all can see (see Tacit and Explicit Information for details). Information only creates value when it interacts with information produced by others. Information that is held by an individual, which is never revealed or acted upon, has no value. The greatest medical discovery in the world does little good if it dies with the discoverer.
- **Knowledge** is the ability to take an action. It is created when information is transformed through human social interactions. A single individual cannot create knowledge. They must interact with information created by others to arrive at an action, a decision. Individuals and organizations must work with the tacit and explicit information generated by others in order to devise a course of action. Often, this course of action is to generate more data, resulting in a new DIK cycle. Knowledge leads to action.
- **Wisdom** encompasses the best, most appropriate action. It usually arises from multiple rounds of the DIK cycle. It requires experience. The DIK cycle often describes an analytical process, one in which simplification is key. Wisdom requires synthesis, often bringing together a wide range of knowledge created from a huge amount of information representing a tremendous mound of data.

8. How Big Data Affects Different Industries

Gartner (2013), Organizations in many industries are collecting and storing a growing amount of data. This is often driven by the need to comply with regulations and/or a mindset that says, "We should keep everything and sort it out later." Although sectors such as government, manufacturing and education have the highest levels of "dark data" (data that is gathered by companies but is not part of their day-to-day operations), it is growing in many industries. Leveraging this underutilized dark data represents the most immediate opportunity to transform businesses.

IT, information management and business leaders can learn from peers in their own industries or, equally importantly, in other sectors. They can use the following collection of research to improve their ability to process growing volumes of data of different varieties, at new speeds. They can then leverage that information to gain an enhanced insight that leads to improved decision making and value creation.

- **"CFOs See the Importance of Business Analytics Improvement";** Gartner Study analyzes the results of the 2012 Gartner Financial Executives International Technology Study. This study shows that improving business analytics is a top technology investment priority for CFOs. IT and information management professionals can use this research to better understand how technologies are prioritized in their organization compared to the benchmark this survey reveals. They can then develop a plan to improve their business analytics capabilities, because their competitors are likely to be doing the same.
- **"A Quick Look at Big Data in Education, 2012"** explains how big data has the potential to inform instructional practice and lead to improved student outcomes. Extracting information from big data and applying it directly in the classroom or lecture hall is a relatively new concept in education, and has a growing momentum at all levels. IT and information management leaders can use this research to start taking advantage of current and emerging solutions to facilitate the translation of data into applicable instructional information.
- **"Best Practices for Big Data Maturity in Financial Services"** explains how banks need to bring big data initiatives under their data governance practices if they want these initiatives to progress and mature. This will require IT and business leaders to re-engineer data governance policies to meet the new data requirements. This research can help financial services organizations use a common definition of big data and work within the enterprise data governance and business intelligence structures to extend data governance practices to encompass the realities of big data.
- **"Target Big Data to Gain Access to Future Strategic Leaders of Healthcare Systems" explores** how leading decision makers in healthcare are targeting strategic excellence through big data analytics. These decision makers are emerging as a robust market for technology providers. As government incentives amplify investments, providers can use Gartner's insight to help them invest in early wins that propel their future success in this market.

9. Key findings from the Survey
In July (2012), Gartner conducted a cross industry survey across 11 vertical industries with the following key findings:
- Education, Healthcare and transportation industries being the early adopters of Big Data technology.
- Around 15-20% of Communication, Media and Services, Insurance, Energy and Utilities industries are planning to invest in Big Data by the end of 2013.
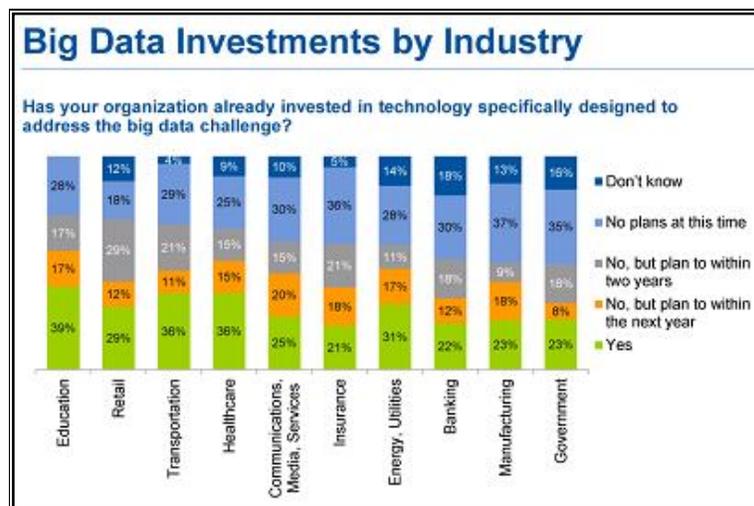- 29% of retail sector have plans to invest in Big Data in next two years.

Figure 3: Big Data Industry Wise Breakup
Source: Gartner (July, 2012)

10. Big Data: Indian Scenario
Some facts about Big Data in India (Ernst & Young Pvt. Ltd., 2013)
- Over the next decade, digital information in India will grow from 40,000 petabytes of data to 2.3 million petabytes, twice as fast as the worldwide rate.
- With 900 million+ mobile connections, 100 million+ active mobile data users and an increasing number of connected devices, the amount of consumer and enterprise data will grow exponentially.
- The diversity of data sources is also rapidly changing with data from social media, the proliferation of video surveillance, audio/images and machine generated data expected to contribute significantly to the data generated.

- According to various industry estimates, the Big Data industry crossed US$8 billion by 2012, and will reach US$16.9 billion by 2015, growing at 7x the ICT industry growth rate.
- As per NASSCOM, the Big Data market in India will grow at 83% annually to reach US$1 billion by 2015.
- Big Data presents an opportunity in the areas of consumer product marketing and innovation, supply chain optimization, smart infrastructure and sustainability management, security and healthcare delivery.
- In India, consumer-oriented industries like BFSI, telecom, retail, healthcare, public sector and manufacturing are among the ones to benefit considerably from Big Data initiatives.

11. Implications for Indian companies (Ernst & Young Pvt. Ltd., 2013)
- The growing importance of unstructured data will require organizations to recognize and tap significant sources of unstructured data.
- Technology vendors can take advantage of proposed investments in scaling up infrastructure capabilities to make unstructured data more important.
- Enterprise wide Big Data adoption activity is likely to remain muted for the next two years as most organizations are yet to start with active pilots.
- IT vendors are still relevant, as high level of outsourcing is expected by organizations to bridge the skill-gap. Vendors should start looking at India as a focused market.
- Big Data related certifications and training is likely to become important. Opportunities exist for companies in the education sector that can design content and impart training to benefit from the demand- supply capability mismatch.
- Vendors can play a meaningful role in driving adoption by conducting workshops, providing ROI calculators, creating analytics lab to demonstrate use-cases that have India relevance. Integrated appliances or solution selling will remain vital for vendors.
- A major factor to ensure the success of Big Data is a large skilled talent pool. In India, Big Data analytics and related IT services will create an estimated 15,000-20,000 specialist jobs by 2015.

12. Key Recommendations for improving Big Data in India (Ernst & Young Pvt. Ltd., 2013)
As organizations work to realize the competitive advantage offered through the adoption of Big Data, Areas of focus that can help CIOs to improve effectiveness of these initiatives:
- Create a "Data centric" culture: CIOs should emphasize on the importance of the data and help organizations in adopting a model where employees can maximize the use of data. Data, when considered as a strategic asset, can help companies in identifying trends previously unseen and opportunities hitherto unknown.
- Invest in building skill sets: The best bet for CIOs is to invest in imparting training to internal staff members (analytics team or BI team) so that they can play the role of Big Data specialists within organizations.
- Measure performance to justify investments: Big Data initiatives should be measured and returns needs to be demonstrable. The internal team, proposing Big Data initiatives, must be able to demonstrate how the technology is going to bring value to the enterprise and how long it will take for these initiatives to start generating returns. The challenge CIOs face is that none of the traditional IT models can be leveraged to calculate the projected return of Big Data investments.

13. Current Global Scenario & Future Outlook

13.1. Current Market Landscape
Indian IT giant Tata Consultancy Services (TCS), 2013 Big Data survey highlights following points related to the current market scenario of Big Data:
- U.S. companies are in the lead
- Per-company spending is a median of $10 million
- 80% of companies have improved business decisions using Big Data
- Unstructured and semi-structured data is now nearly half the total data
- Few companies sell their data, even though the revenue is substantial
- Companies expecting ROI from Big Data anticipate high returns
- Digital businesses better harness the power of Big Data
- Sharing data across organizational silos is the biggest challenge

Further the Open source analyst firm estimates the current Big Data market at just over $5 billion in total factory revenue. The market is now becoming increasingly competitive with IT majors such as IBM & HP along with a traditional pure play like Vertica and Cloudera are bringing in significant revenue today helping enterprises, governments and healthcare organizations, process and make sense of the torrents of unstructured data flowing from mobile devices, sensors, social media and other sources with the business of Big Data is only going to grow as the volume of data increase.

14. Future Outlook

The Big Data market is emerging at a rapid pace and incorporating technology and services from a wide range of existing and new market segments. Although there are multiple scenarios that could unfold and many demand and supply variables remain in flux, one can expect the market to exhibit strong growth over the next five years. Some of the major forecast would include (TCS, 2013):

- **North America will continue to lead in Big Data:** North America will continue to be the leading market for Big Data services and products with a median spend of about $10 million. Europe comes next; however will be slower to realize gains from Big Data initiatives due to policies and regulations. Growing trends indicate that US, India, UK and Mexican companies are likely to invest in Big Data initiatives in the coming years.
- **North American companies also have better vision regarding data assets:** Asian countries generate more unstructured data, often multi-lingual, thus making unstructured data analytics such as text mining, sentiment analysis, and content analytics more challenging. Until unstructured data processing and analysis become more widespread and mature, structured data assets will lead the way and thus North American companies will continue to have the edge.
- **Data Assets will become the basis of competition and valuation:** Companies such as Facebook, Twitter and Netflix are now being viewed beyond their products and services they offer; the collected data through interactions and transactions from millions of users will soon become a basis for company valuation. Companies will start to differentiate themselves in both the quantity and quality of data assets they have for internal and external use.
- **Big Data spending on products and service will increase through 2020:** In coming years, leaders will be quicker to realize the business value and will put things in place to implement Big Data projects in an agile way. Many enterprises will be driven to consider Big Data processing due to competitive pressures. This indicates that business process optimization would be needed as a result of Data analytics.
- **The need for speed:** Business users are beginning to realize that data analysis at scale is becoming important for competitive differentiation; Business users will demand even faster analytical results and will want answers as quickly as possible. Key stakeholders and business users will start to view Big Data initiatives with a key metric in mind: the Time-to-Insight metric.

Further research firm International Data Corporation (IDC's) estimate key trends and forecast assumptions that will impact the Big Data market in the foreseeable future include:

- Information overload and the high cost of finding the right information are significant issues for many organizations and present an opportunity that can be addressed with Big Data technology.
- Global Big Data market revenues are forecast to reach $12.4 billion in 2013. A key driver of Big Data services is their applicability to enterprises where some type of data is compiled and managed. Enterprises in the healthcare, communications, retail and transport sectors are adopting Big Data services more rapidly than other sectors.
- IDC believes Big Data market will increase at a 31.7 percent compound annual growth rate (CAGR), reaching $23.8 billion in 2016. This CAGR is seven times faster than the growth rate expected for the overall information and communication technology market, IDC noted.
- Big Data services are expected to grow at a 21.1 percent CAGR, while Big Data storage is forecast to grow at a more aggressive 53.4 percent CAGR due to the current dual use of storage in Big Data environments
- A shortage of Big Data technology skills may fuel demand for cloud-based Big Data products, resulting in an increasing focus on Big Data automation and analytics.
- The segment of the Big Data market that is based on open source technologies will develop along the lines established by the development of the Linux ecosystem, where much of the revenue will be accrued by hardware and applications and application development and deployment software vendors.
- There are a large number of open source projects offering Big Data technology that does largely the same thing. Within a few years, most of these will be relegated to the category of obscure technologies for specialized practitioners as commercial vendors take over the Big Data market. Others will become casualties of large IT vendors' entry into the market either through their own R&D or acquisitions.
- Big Data appliances that include integrated software and hardware will expand rapidly, partly as a response to alleviate the deficit in expert IT skills associated with optimizing hardware and software to work together for Big Data workloads. These Big Data appliances will be deployed both on-premises and in the cloud.
- Buyers of Big Data technology will include not only end users, but also service providers that will in turn provide value-added services — often industry-specific services — to end users.

15. Conclusion

Big Data will become a means of value creation for businesses, tapping into the untamed and the vague world of information. As the quantity and interconnectedness of data immensely increases, the value of the Big Data will foster. Big Data can unlock significant value by making information transparent and usable at much higher frequency. Through better analysis of large volumes of data that are becoming available, there is the potential for making faster advances in many scientific disciplines and improving the profitability and success of many enterprises. Big Data has proved to create value at almost all fronts be it organizations, customers, workforce, or society as a whole. For organizations, sophisticated analytics with adoption of Big Data will substantially improve decision-making leading to increased ROI. Effective use of Big Data techniques can improve the

development of the next generation of products and services. For instance, manufacturers using data obtained from sensors embedded in products to create innovative after-sales service offerings such as proactive maintenance. Right information at the right time and the right place can be obtained through Big Data, which would prove essential to organizations sustainable development and competitive advantage; however, one of the key challenges faced by organizations would be lack of knowledgeable, skilled and trained manpower. Big Data is the need of the hour and is going to stay. Big Data has grown in the last decade and current trends shows that organizations will be adopting this technology in times to come. According to Markets and Markets the global Big Data market will reach $46.34 billion by 2018, which ultimately means large no. of job opportunities shall be created at this front. Big Data will create jobs in three areas: Data Infrastructure, Data Management, and Data Presentation (visualization). Large data sets need storage, Database software to store the data, and Analytics tools to turn data into meaningful information for businesses. Additionally, there are other parts of IT that will be impacted by the surge of Big Data adoption such as Network, Security, and the Cloud. Major Universities and B-Schools in India should provide special internationally renowned courses, degree or diploma, in Data Tapping Techniques like Big Data. Manpower should be created to capture this new technique so that India can have an edge.

16. References

1. 2013 IT Market Trends and Predictions on Big Data. Retrieved from http://blogs.sap.com/innovation/big-data/2013-it-market-trends-and-predictions-big-data-028908
2. Commonwealth of Australia. (2013). Big Data Strategy — Issues Paper. Retrieved from http://agimo.gov.au/files/2013/03/Big-Data-Strategy-Issues-Paper1.pdf
3. CRISIL. (2013). "Big Data: The Next Big Thing," website. Retrieved from www.crisil.com/global-offshoring/granasscom.html.
4. Ernst & Young Pvt. Ltd. (2013). "Big Data and Enterprise Mobility Growing relevance of emerging technology themes: the India perspective". Retrieved from http://www.ey.com/Publication/vwLUAssets/Big_Data_and_Enterprise_Mobility/$FILE/Big_Data_Enterprise_Mobility_LR.pdf
5. Gartner (2012). Invest in Information and Analytics to Benefit From Big Data. Retrieved from http://www.gartner.com/id=2363115
6. Gartner. (2011). Pattern-Based Strategy: Getting Value from Big Data. Retrieved from http://www.gartner.com/it/page.jsp?id=1731916
7. Gartner. (2012).Gartner reveals Top Predictions for IT Organisations and Users for 2013 and beyond. Retrieved from http://www.gartner.com/it/page.jsp?id=2211115
8. Gayle, Richard. "The DIKW Model for Innovation" Retrieved from http://www.spreadingscience.com/pdfs/1_DIKW.pdf
9. IBM Research. (March 2012). "2012 IBM Global Technology Outlook". Retrieved from http://www.research.ibm.com/files/pdfs/gto_booklet_executive_review_ march_12.pdf
10. IDC. (2013). "Digital Universe study". Retrieved from wwwindiatechonline.com/it-happened-in-india.php?id=323,
11. IDC. (March 2012). Big Data: Global Overview: Market Analysis Retrieved from www.idc.com/getdoc.jsp?containerId=prUS23355112
12. Kelly.J. (2013). "Big Data Vendor Revenue and Market Forecast 2012-2017". Retrieved from http://wikibon.org/wiki/v/Big_Data
13. McKinsey Global Institute, (June 2011). "Big Data: The Next frontier for innovation, competition and productivity". Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
14. TCS. (2013). "The Emerging Big Returns on Big Data, 2013 Global Trend Study". Retrieved from http://www.tcs.com/big-data-study/Pages/default.aspx
15. TWDI. (2011). "Big Data Analytics". Retrieved from http://tdwi.org/research/2011/12/sas_best-practices-report-q4-big-data-analytics.aspx?tc=page0