Leveraging Big Data for Enhanced Business Values

Ravi Kumar
Module Leader,
Persistent System Ltd,
Nagpur

Rashmi Bhadani
Visiting Assistant Professor, Prestige Institute of Management and Research
Indore

ABSTRACT
In today's age of globalization and internalization, precise and meaningful data plays a fundamental role in making crucial decisions across every strata of business, whether it is associated with human resource or sales or production or marketing or for any other field. For example, Apple, Facebook, Twitters and others know the next emerging trend in the market and transmute its strategies to adapt to the recent trends. All these multi-billionaire giants have the power of analysing the myriad of data flowing in from sundry of sources like social media, market trends, past annals and lot more. They are all accoutred with the most-indispensable tool in current globalized economy, "The BIG DATA." It refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse. With the advent of Big Data, the business houses now can capture various trends in the market with minimal cost incurred. It had helped individuals as well as the organization to make proper decisions and increase the base of their operations by targeting the apt customers for the business. Big Data is playing a pivotal role in transfiguring the business decisions. On the other side, there are numerous challenges presented by the 'Big Data," like security and privacy issues, infrastructure failure, and others. Thus, this paper presents the concept of Big Data, opportunities derived by companies using cases, and challenges related to Big Data. Since, Big Data is in its embryological phase; this paper will provide a snapshot to the scholar for future research in this field.

Keywords
Big Data, privacy, security, opportunities

1. INTRODUCTION
Ubiquitous Data Torrent: Data, Data Everywhere. For years long, large amount of data is needed for generating valuable business insights. Earlier, the census method was used to record data that were related to crop yields. Later, analysis of these collected data was done for improving the effectiveness of society. Not only limited to this, the foundational theories in calculus, statistics, and probability fostered new tools that were used by scientist to predict movements of stars and suns, predicting rates of various crimes, event of suicides, and others. One such important development was related to the field of collection and analysis of data was done by Dr John in 1800's. Dr John has used modern tool of data analysis to map the cholera clusters across London and thus laid the foundation of the germ theory.

In today's world of technological revolution and networking, data has been woven in our life as threads are woven in fabrics. This huge amount of data generation is from diverse sources, like mainframe, mobile devices, interaction between client and server, social networking, and others. Instead of discussing the entire advent of data conglomeration and multifarious data growth, focus can be concentrated on a time frame of just 60 seconds for consideration. For every 60 seconds, in this globalized and digitized world, the volume, velocity and variety of social data and mobile data can be gauged from the following statistics [1]:

- 694,445 queries are served by Google search engine.
- More than 6,600 images are uploaded on website named Flickr.
- More than 600 videos that account for more than twenty-five hours contents are uploaded on the YouTube.
- More than 168,000,000 electronic mails are sent on the network.
- Creation of 320 fresh accounts on Twitter.
- 98,000 new tweets on Twitter, a social networking websites.
- Downloading of iPhone application around 13,000 times.
- 695,000 updates of status, 510,040 comments, and wall posts comprising of 79,364 postings on Facebook, and others.

Additionally, CISCO estimates the growth of annual data traffic on the network to 7.7 zettabytes in 2017 [2]. In a nutshell compendium, the today's data is huge (Volume), multi-structured (Variety) and fast Moving (Velocity). This further strengthens the fact that an insight into data will help not only in staying competitive, but also reaching the right customer base, at right time, for the right business. Further, traditional data systems are unable to handle the data or analytical needs of the present era. Primary reason can be contributed to the amalgamation of structured and unstructured data and the time required for retrieving the desired information for the substantial data. Present data stack is stringent to the changes required for the meticulous analysis of data for obtaining a useful insight having economic value. These scenarios provided the necessary impetus and impelled the creation of a new data stack.

2. BIG DATA CONCEPT
The new data pile consisting of Big Data is the panacea for overcoming all the hurdles of traditional data sets. Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to

be in order to be considered big data—i.e., we don't define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes) [3]. Gartner's defines Big Data as high volume, -variety, -velocity information assets which demands innovative way of processing and cost effective for enhancing the decision-making and insights of business [4]. IDC defines the term big data as one of the new generation invention in the field of technologies as well as architectures that are designed to extract value economically from very large volumes diversified data through enabling high-velocity apprehension, discovery, and analysis. This incorporates hardware as well as software that integrates, organizes, manages, analyzes, and presents data that is characterized by the "four V's": Volume, Variety, Velocity and Value [5].

With more and more people and business, joining the digital revolution, the amount of data being generated had grown exponentially. The quantum of data produced had raised from hundreds of terabytes to thousands of petabytes data and is still on the rise. With big data in place, useful insight from the data can be generated. For example, database for yahoo Search assist Tm is based on Hadoop. Hadoop forms the base of big data technology. Big data in these days have been the nave of Yahoo's business [6]. Centered on the big data technologies, Yahoo have developed an entire model of revenue as well as monetized the data that has been collected. Yahoo advertising strategy helps advertisers to reach appropriate base of customers.

The lifestyle of modern day is revolving around social networking websites like Facebook, Twitter, Google, and others. Facebook had leveraged the Big Data platform to embellish its services like Facebook Messaging, as well as Facebook Insights. Facebook Insights fosters developers as well as website owners with accessibility for real-time analysis that is related to Facebook activity throughout websites with social plugins, Facebook Ads, Pages, and others. Using data, Facebook surfaces activities such as ions, website visits, and click through rates. Thus, the said analysis could foster individual, businesses, bloggers, and other to gain valuable insights into the way people are interacting with the content as well as optimize its services [7].

3. LEVERAGING BIG DATA: CASES

As per report published by McKinsey Global Institute of Analysis in 2011, Big Data has the potential to generate significant value across sectors. It has been seen that Big Data fosters integrate storage, application, as well as analytics. It helps in achieving great business values, like customer oriented marketing, enhanced efficiency, increased quality, analysis of customer activity, personalized products as well as services, and increased level of customer experience and satisfaction. Some of the company's that has reaped the benefits of "Big Data" are:

3.1 Case1: Cisco

Cisco had the challenge to unlock the business value of large datasets, which comprises of both structured, as well as unstructured information. It implemented the enterprise Hadoop platform for big data and provided a new incarnation to its business model. The model was capable of finding promising opportunities related with the sale. As a result of Big Data, CISCO recorded $40 million in incremental booking for the fiscal year 2013-2014. Srini Nagpuri, CISCO IT Project Manager, shed some light on Big Data platform, which is in place for their company. It processes 1.5 billion records daily, and it discerned opportunities on the very first day of its deployment in the production.

3.2 Case 2: Boeing

Boeing is the forerunner in the field of aviation and defence. It also provides copious of analytics software to distil information out of IP traffic and flow data in real time. This business model leverages the real-time data and mammoth-amount of data must be processed in a jiffy. With the help Hadoop's contrivance, they were able to slash the expenses of maintaining the servers by half. They incurred a profit of more than $250,000 during a span of three years since the inception of the Big Data project [10].

3.3 Case 3: General Electrical (GE)

With the advent of 2011, the Fairfield, Connecticut-based industrial giant GE announced a $1 billion investment to build software and expertise for its version of Big Data analytics. GE is a $145 billion company engaged in the manufacturing of jet engines, power plants, and locomotives. They wanted to monetize the profusion of data disgorging out from proliferation of devices manufactured by the company. These include data from, jet engines, turbines, and hospital MRI equipment. Harvesting such data, will enable the company to isolate maintenance problems, improving fuel efficiency and make other operational improvements. William Ruh, vice president and corporate officer of the big data analytics center, explains about empowering machine with intelligence and the need of delivering right data to the right person at right time. GE aspires to make $30 trillion with the help of current investment by 2030 [11].

4. BIG DATA: CHALLENGES

Some of the challenges due to the inherent characteristics of Big Data are:

4.1 Issue with validity of data

The volume feature of Big Data also presents a huge challenge for the organization as well as individuals. It is because large quantity of data requires continuous cleaning up process. Spending much of the time in the information cleaning process will lead to the production of information that is too late given to the users or is no longer useful.

4.2 Privacy and security violation

Every individual, as well as organizations, have the right to protect their privacy. Violation of privacy takes place when individual's credentials stored in their online profiles are used by the companies for marketing purposes. Organizations also suffer the risk of losing their sensitive data related to financial information or client information or others. Security of data is a threat when data is hijacked by stealing the confidential credential and other sensitive user data; or manipulation of data; or others [8].

4.3 Infrastructural fault

Infrastructural fault can lay all benefits derived from Big Data into vain. Considering the large amount of complex data, there is a requirement for reliable hardware and software system. For instance, server failure when uploading of a file is going on. This infrastructure fault can also be in case where hard disk or other computing resources halts during the process of data transmission.

4.4 Lack of talented human capital

The field of big data is still in its pilot phase. In addition to this, there is shortage of skilled IT professionals for analyzing as well as using these data strategically.

5. DISCUSSION

Big Data presents numerous opportunities for the individual's as well as organizations to enhance their sustainability. At the same time, it is necessary to focus on certain challenges raised due to Big Data. For instance, reaping benefits of big data requires huge computing power of the hardware, software, and stable networks. It can be achieved through cloud computing which offers rapid scalability, shared pool of resources, and pay-as-per use. Cloud computing will help individuals as well as organizations to speed up the data processing operation. Further, the problem of large volume of data and delayed information retrieving can be resolved through implementing automatic data filtering system. This automatic data filtering system will transfer the crucial information to the data clusters for immediate analysis as well as led crucial data is stored at different clusters that can be later analysed or removed. One of the most widely used technologies for reducing as well as sorting data is Hadoop. The challenge related to privacy violation can be resolved through encrypting the data as it has been said that the confidentiality of data can be protected through encrypting it [9]. Strategic management should be done to store backup of the data if any catastrophic event takes place. This is necessary as no individuals or companies can afford to lose their valuable data.

6. CONCLUSION

In a world of cut to cut competition, reaching the right base of customers and leveraging the current IT setup for enhancing productivity, has become a paramount concern for any business. Big Data rightly fits the above criteria and is the new-buzzing word in the business houses. It refers to data set which is large, diversified, and complex. Gleaning insights from data had helped to gain deeper insight, boost the business, increase economic activity, and gain competitive advantage. It is playing a pivotal role in moulding the new dynamics of business, expediting data processing and accruing profits from annals of the past and present. Thus, it can be plausibly said that Big Data is transfiguring business and adding a new dimension to decision making.

7. REFERENCES

[1] Kalakota Ravi. 2012. Sizing Mobile + Social Big Data Stats. Retrieved from http://practicalanalytics.wordpress.com/2012/10/22/sizing-mobile-social-big-data-stats/

[2] Cisco. 2013. Cisco Global Cloud Index: Forecast and Methodology, 2012-2017. Retrieved from http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html

[3] Manyika James, Chui Michael, Brown Brad, Bughin Jacques, Dobbs Richard, Roxburgh Charles, Byers Hung Angela. 2011. Big Data, The next Frontier for Innovation, Competition and productivity. McKinsey Company, Published in May 2011

[4] Sicular Svetlena. 2013. Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s", Published on March 27, 2013. Retrieved from http://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/

[5] Srirangpatna Dinesh, Dialani Mukesh, and Muscolino Holly. 2014. The Big Data Analytics Services Opportunity for the Hardcopy Industry. IDC: Industry Development Models.

[6] Radia Sanjay. Data Processing in the Cloud at Yahoo!, Retrieved from http://www.cs.bris.ac.uk/~simonm/conferences/isc09/Radia.pdf

[7] Borthakur Dhruba et. al. 2011. Apache Hadoop goes Real time at Facebook. SIGMOD'11 June 12-16, Athens: Greece.

[8] CSA: Cloud Security Alliance. 2010. Top threats to cloud computing. Cloud Security Alliance, March 2010.

[9] Zhang, Qi., Cheng, Lu, and Boutaba Raouf. 2010. Cloud computing: state-of-the-art and research challenges. Journal of Internet Services and Applications, 1(1):7–18, May 2010. doi: 10.1007/s13174-010-0007-6.

[10] MAPR Technologies. 2012. Boeing: Augmenting Real Time Classification for speed, cost saving and ease to use. Retrieved from http://www.mapr.com/sites/default/files/mapr_boeing_narus_sb0113.pdf

[11] TCS. 2012. Big Data Case Study: How GE is Building Big Data, Software and Analytics Capabilities for an "Industrial Internet".