SGI-JOURNAL OF SOCIAL SCIENCE AND HUMANITIES
VOLUME. 1
ISSUE. 4
AUGUST, 2020

GROWTH TOWARDS BIG DATA ANALYTICS IN COMPETITIVE MARKET

[1] **Venkerta N. PhD**.

[1] Panimalar Engineering College, Department of Management Studies, Chennai.

Abstract

Big Data represents a fundamental shift in business decision making. Organizations are used to analyzing internal data– sales, shipments and inventory. Now they are increasingly analyzing external data too, gaining new insights into customers, markets, supply chains and operations.

The need for Big Data solutions is becoming a reality for more and more companies. This paper provides valuable insight into developer opinions. Over the last 20-30 years with the implementation of enterprise applications, there is a tremendous amount of data that has been captured. The early wave of business intelligence tools have attempted to provide reporting and analysis of this data.

The rapid growth of cloud computing and the Internet of Things (IoT) further promote the sharp growth of data. In the paradigm of IoT, sensors all over the world are collecting and transmitting data which will be stored and processed in the cloud. The intent of this paper is to provide supplemental guidance about the business factors that drive the adoption of big data analytics and also by explaining the reasons why adoption is compelling.

Up until the mid-nineties, data acquisition and processing have primarily been carried out manually using passive keyboards. The article also focused that appropriate investment in Big Data will lead to a new wave of fundamental technological advances that will be embodied in the next generations of Big Data management and analysis platforms, products, and systems.

Big data allows manufacturing, production and research processes to be more efficient and cost-effective. It should help cut through the clutter of opinion and guide a targeted approached based on underlying insights. Those who can successfully distill the data will have a significant competitive advantage.

More people have more access to more technology than ever before. This article shows unmistakable signs that technologies used to manage, analyze, report and make business decisions from big data are becoming easier to use and are more widely available to employees in companies large and small.

Keywords: *Big Data, Data Storage, Hadoop Distributed File System (HDFS).*

I. Introduction

The theme of Big Data has spawned a tremendous amount of attention and investor interest in recent years. Companies are at varying stages of adoption of Big Data analytics and as with all new technologies the maturity curve of all firms are emerging that delineates the early adopters from those who are at a more foundational level. By introducing analytics and more flexible production techniques, manufacturers, for instance, could boost their productivity by as much as 30 percent.

Much data today is not natively in structured format; for example, tweets and blogs are weakly structured pieces of text, while images and video are structured for storage and display, but not for semantic content and search: transforming such content into a structured format for later analysis is a major challenge.

McKinsey (2011) projected in an industry report that five new kinds of value might come from big data: (1) creating transparency in organizational activities that can be used to increase efficiency; (2) enabling more thorough analyses of employee and system performance in ways that allow experiments and feedback; 8(3) segmenting populations in order to customize actions; (4) replacing/supporting human decision making with automated algorithms; and (5) innovating new business models, products, and services (McKinsey Global Institute, 2011).

We are awash in a flood of data today. In a broad range of application areas, data is being collected at unprecedented scale. Decisions that previously were based on guesswork, or on painstakingly constructed models of reality, can now be made based on the data itself. Such Big Data analysis now drives nearly every aspect of our modern society, including mobile services, retail, manufacturing, financial services, life sciences, and physical sciences.

Big data is huge volumes of fast moving, structured and unstructured data in petabytes (1,024 terabytes) or exabytes (1,024 petabytes), which is almost impossible to process and make sense of using traditional databases. Such data, sourced from multiple channels like internet, social media or mobile data, if rightly assimilated and interpreted, opens up a plethora of opportunities for the businesses be it understanding consumption pattern or industry trends and trump competition.

Big data analytics can help cope with large data volumes, data velocity and data variety. Enterprises have started leveraging these big data systems to mine hidden insights from data. Big data analytics can make a significant impact on several industry verticals like medical, retail, IT and how enterprises can harness the value of big data.

Big data enables quantification and analysis of human life, from the macro down to the micro level—that is, from the global scale down to the individual. It will present new opportunities for governments, businesses, citizens, and organizations. Big data could be used to help solve or manage critical global problems, lead to new scientific breakthroughs and advances in human health, provide real-time information and analysis on wide areas of life, wire up the planet's natural systems for monitoring and environmental remediation, greatly improve efficiency and resource use, and enhance the decision-making and daily operations of society.

According to Intel, each internet minute sees 100,000 tweets, 277,000 facebook logins, 204-million email exchanges, and more than 2 million search queries fired. Looking at the scale at which data is getting churned it is beyond the scope of a human's capability to process data and hence there is need for machine processing of information. In the recent past, the industry has grown significantly – by almost 14 percent in 2012 – and is slated to grow to a size of $50.7 billion by 2016. It is being forecasted that the market for analytics and business intelligence (BI) platforms will be the fastest growing segment in the software markets. With emerging trends such as data-as-a-service coming, analytics shall probably see further growth prospects.

Today, the creation and use of big data expand beyond large web companies like Yahoo, Google, and Face book. Businesses everywhere, including industrial enterprises, face mounting pressure to stay competitive with data-driven strategies—requiring increasingly more data, which results in the accumulation of larger and larger data sets. In addition, evolving and evermore stringent regulatory requirements necessitate the collection of more information as proof for audit and compliance purposes.

Today most data is generated through machine to machine communication through contactless technologies such as radio frequency enabled devices (RFID), mobile phones and other embedded devices. Three-quarters of decision-makers (76 per cent) surveyed anticipate significant impacts in the domain of storage systems as a result of the "Big Data" phenomenon.

Analytics is usually a dominant process in industries/ domains that are highly data-rich and the major importance of its usage is attached to the reduction of business risks, improvements in revenue accrual, and generally in increasing overall operating efficiencies. As per available records, data and analytics functions are most popularly used in areas of sales and marketing, followed by customer service and R&D, and peripherally by IT and manufacturing.

With the help of current developments, companies can combine with new external data sources from all industries advanced analytics. With the technologies like Hadoop, in-memory computing and Cloud Infrastructure-as-a-service, analytics and visualization is entering new phase of consolidation. Hadoop enables data storage scale through the use of commodity hardware, distributing data across many low-cost computers.

II. Literature Review

Recognizing the importance of big-data, the White House formed the Big Data Senior Steering Group in 2010, a joint effort of 17 agencies, to identify current Federal big data research and development activities (The White House, 2012).

In 2010, Apache Hadoop defined big data as "dataset wh ich could not be captured, managed, and processed by general computers within an acceptable scope". On the basis of this definition, in May 2001, McKinsey & Company, a global consulting agency announced Big Data as "the Next Frontier for Innovation, Competition, and Productivity."

In addition, the US National Institute of Standards and Technology (NIST) defines big data as "Big Data shall mean the data of which the data volume, acquisition speed, or data representation limits the capacity of using traditional relational methods to conduct effective analysis or the data which may be effectively processed zoom technologies," which focuses on the technological aspect of big data.

Scientific research has been revolutionized by Big Data [1]. The Sloan Digital Sky Survey [8] has today become a central resource for astronomers the world over. The field of Astronomy is being transformed from one where taking pictures of the sky was a large part of an astronomer's job to one where the pictures are all in a database already and the astronomer's task is to find interesting objects and phenomena in the database. In the biological sciences, there is now a well-established tradition of depositing scientific data into a public repository, and also of creating public databases for use by other scientists. In fact, there is an entire discipline of bioinformatics that is largely devoted to the curation and analysis of such data. As technology advances, particularly with the advent of Next Generation Sequencing, the size and number of experimental data sets available is increasing exponentially.

A.T. Kearney forecasts global spending on Big Data hardware, software and services will grow at a CAGR of 30% through 2018, reaching a total market size of $114B. The average business expects to spend $8M on big data-related initiatives this year. Source: Beyond Big: The Analytically Powered Organization.It is widely believed that the use of information technology can reduce the cost of healthcare while improving its quality [2], by making care more preventive and personalized and basing it on more extensive (home-based) continuous monitoring. McKinsey estimates [4] a savings of 300 billion dollars every year in the US alone.

Not surprisingly, the recent PCAST report on Networking and ITR&D [5] identified Big Data as a "research frontier" that can "accelerate progress across a broad range of priorities." Even popular news media now appreciates the value of Big Data as evidenced by coverage in the Economist [1a], the New York Times [7], and National Public Radio [6].

Industry analysis companies like to point out that there are challenges not just in Volume, but also in Variety and Velocity [3], and that companies should not focus on just the first of these. By Variety, they usually mean10 heterogeneity of data types, representation, and semantic interpretation. By Velocity, they mean both the rate at which data arrive and the time in which it must be acted upon. While these three are important, this short list fails to include additional important requirements such as privacy and usability.

Challenges of Big Data

Today data management and analytics systems are based on the relational database management system (RDBMS). In addition, RDBMS, are increasingly utilizing more and more expensive hardware. For solutions of permanent storage and management of large scale disordered datasets, distributed file systems and NoSQL databases are good choices. Various big data applications can be developed based on these innovate technologies or platforms.

Some of the key challenges of Big Data which include: Data Representation, Redundancy Reduction and Data Compression, Data Life Cycle Management, Analytical Mechanism, Data Confidentiality, Energy Management, Expendability and Scalability and Cooperation.

Growing Digital Exposure on Human Interaction

Most of the human ingenuity across globally to solve problems is through crowd-sourcing. Wikipedia, the online encyclopedia, is perhaps the best known example of crowd-sourced data. We are relying upon information provided by unvetted strangers. Most often, what they say is correct. However, we should expect there to be individuals who have other

motives and abilities – some may have a reason to provide false information in an intentional attempt to mislead.

Every day we create 2.5 quintillion bytes of data - so much that 90% of the data in the world today has been created in the last two years alone. Similar to the complexity aspect of Big Data, its growth rate is mostly due to the ubiquitous nature of real-time data capture devices, systems and networks. We can expect that this growth rate will continue to increase in the future.

Leaders in every sector will have to grapple with the implications of big data, not just a few data-oriented managers. The increasing volume and detail of information captured by enterprises, the rise of multimedia, social media, and the Internet of Things will fuel exponential growth in data for the foreseeable future.

Every person with a mobile phone can act as a multi-modal sensor collecting various types of data instantaneously (e.g., picture, video, audio, location, time, speed, direction, acceleration).

Businesses know that measurement is one of the keys to management. Collecting and understanding data about how an organization operates leads to knowledge that can improve decision making, refine goals and focus efforts. Providing quality data in the right format is becoming an increasingly important factor. The worlds of data collection and analysis, sophisticated business software applications, and accepted measurement standards are coalescing to help drive transparent and improved sustainability performance for companies and their supply chains.

We are currently working and living in digital universe which is made up of images and videos on mobile phones uploaded to YouTube, digital movies populating the pixels of our high-definition TVs, banking data swiped in an ATM, security footage at airports and major events such as the Olympic Games, subatomic collisions, transponders recording highway tolls, voice calls zipping through digital phone lines, and texting as a widespread means of the Internet of Things (IoT) helps digitize factories, making humans fitter to meet the challenges of the 21st century economy (international competition, increasingly shorter production and innovation cycles, growing demand for customized goods, etc.).

Big data analytics require new techniques and algorithms that are yet to be developed, requiring new research and development. They will also, in time, benefit from the near-real-time analytical abilities of quantum computing. Data have swept into every industry and business function and are now an important factor of production, alongside labor and capital.

Big data provides massive and exponentially growing digital traces of human activity that is becoming a laboratory for understanding society. All kinds of data—videos, blogs, Internet searches, social media postings, tweets, and so forth—can be analyzed to better understand and predict human behavior.

While the portion of the digital universe that holds potential analytic value, including economic value, is growing—from about 25 percent today to about one third of an exponentially larger universe in 2020—only a very small portion has so far been explored. There are large untapped pools of data in the digital universe from which big data technology can potentially extract value.

From a recent study it was concluded that 72% of companies are using social technologies in one way or another, while very few are close to taking full advantage of implementing the technology. Though the most powerful applications of social tools worldwide remain largely unused, enterprises are on their way to developing ways to reach customers via social technologies and gather insights for product development, marketing and customer service.

For individuals, big data holds the potential of a huge new universe of useful information to enhance the quality of life, health care, and improved services of all kinds, including government services. But individuals will also be potentially haunted by their “digital exhaust”

that will in some instances last longer than their lifetimes. The use of big data will become a key basis of competition and growth for the private sector, as it has been for Amazon, Google, and others, thus impacting overall economic growth.

Growth Rate of Big Data

Large enterprises are focusing on Big Data projects as more companies move from small, proof-of-concept projects to large-scale, production-level deployments. A quarter of decision-makers surveyed predict that data volumes in their companies will rise by more than 60 per cent by the end of 2014, with the average of all respondents anticipating a growth of no less than 42 per cent.

Big Data technologies are going to get commoditized in the next couple of years. New technologies like Hadoop, HBase etc will mature with both their skills and partner ecosystem getting more diverse and stable. Increasing number of vendors will offer very similar capabilities and we will see these vendors compete increasingly on operational efficiency on the pivots of speed and cost.

Data production will be 44 times greater in 2020 than it was in 2009. According to Twitter's own research in early 2012, it sees roughly 175 million tweets every day, and has more than 465 million accounts.

The global market for Big Data was projected to grow 61% from the prior year to $18.1 billion in 2013. Stagnant growth is forecast for the overall global economy through 2017, but analysts expect the Big Data market to continue growing quickly. This growth will stem from an increased recognition by enterprises that Big Data analytics can help them "do more with less".

Big data is a top business priority and drives enormous opportunity for business improvement. Wikibon's own study projects that big data will be a $50 billion business by 2017. As recently as 2009 there were only a handful of big data projects and total industry revenues were under $100 million. By the end of 2012 more than 90 percent of the Fortune 500 will likely have at least some big data initiatives under way.

In 2014, $15 billion is expected to be spent on big data technology compared with less than $10 billion two years ago. Spending growth in the big data space is expected to be an annualized 27 percent over the next five years, compared with 4.5 percent growth for overall IT spending. New software tools, like Microsoft's Azure Intelligent Systems, are helping to fuel growth by streamlining data analysis processes.

III. Key Findings

- *Zetabytes of data exist in the digital universe today.*
- *The Obama administration is investing $200 million in big data research projects.*
- *Global big data market is estimated to be $14.87 billion in 2013 and expected to grow to $46.34 billion. This represents an estimated Compounded Annual Growth Rate (CAGR) of 25.52% from 2013 to 2018.*
- *By 2018 big data requirements will gradually evolve from differentiation to 'table stakes' in information management practices and technology*
- *Facebook stores, accesses, and analyzes 30+ Petabytes of user generated data.*
- *Walmart handles more than 1 million customer transactions every hour, which is imported into databases estimated to contain more than 2.5 petabytes of data.*
- *More than 5 billion people are calling, texting, tweeting and browsing on mobile phones worldwide.*
- *Akamai analyzes 75 million events per day to better target advertisements.*
- *94% of Hadoop users perform analytics on large volumes of data not possible before; 88% analyze data in greater detail; while 82% can now retain more of their data*

Future Directions

In the future huge scientific projects, winning game-shows, and medical research fit in naturally especially with big data analytics and decisions will be based on data and analysis rather than experience and intuition that is the bigger dimension of the Big Data phenomena. By 2020, big data features and functionality will be non-differentiating and routinely expected from traditional enterprise vendors and part of their product offerings.

Perhaps Big Data is a way of defining a paradigm shift to a data-intensive collaboration where processes reinforce traditional database approaches. There is a considerable amount of hype which will ultimately lead to a degree of disillusionment.

IV. Conclusion

Business and IT leaders need to ask themselves whether their industrial enterprise is maximizing the full potential value of their process data and using that insight to drive real-time improvements. As data volumes continue to expand, information-driven strategies will only become more pervasive as a source of competitiveness—making the use of big data in the industrial space ever more imperative.

Successful enterprises will seize this opportunity by advancing their cultures to drive an enterprise-wide analytic operating model that puts data science at the center of their business strategy and management practices.

Big data and analytic practices represent a new Industrial Revolution, and these areas will continue to evolve in exciting new ways. [9] This is a transformational moment for enterprises, as business analytics presents both opportunities and challenges.

Through better analysis of the 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.

Going forward, as information increasingly empowers enterprises to understand their businesses better and to foresee what is possible, those that capitalize on the value of big data will gain insights to improve performance beyond their competitors. They will be positioned to better innovate, compete, and drive value—all of which will significantly accelerate business growth and continuously drive optimized performance for long-term success. Overall, Big Data has the potential to create the next wave of successful technology companies that could change the way we all live and work.

References

The White House, Office of Science and Technology Policy. (March 29, 2012). Big data is a big deal. Advancing Discovery in Science and Engineering, Computing Community Consortium Spring 2011.

Drowning in numbers -- Digital data will flood the planet — and help us understand it better. The Economist, Nov 18, 2011 Smart Health and Wellbeing Computing Community Consortium Spring 2011.

Pattern – Based Strategy: Getting Value from Big Data Gartner Group press release July 2011.

Big data: The next frontier for innovation, competition, and productivity. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. McKinsey Global Institute May 2011.

Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology PCAST Report, Dec. 2010.

Following the Breadcrumbs to Big Data Gold Yuki Noguchi. National Public Radio, Nov. 29, 2011.

The Age of Big Data. Steve Lohr. New York Times, Feb 11, 2012.

SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way Galaxy, and Extra-Solar Planetary Systems. Jan. 2008.

The Science of Big Data, an interview with Fermilab physicist Rob Roser, at www.atkearney.com.