

# Big Data idea implementation in organizations: potential, roadblocks

Nishchitha S<sup>1</sup>, Anusha N M<sup>2</sup>, Priya D S<sup>3</sup>

Nishchitha S, Saphthagiri College of Engineering, Bangalore, Karnataka, India

Anusha N M, Saphthagiri College of Engineering, Bangalore, Karnataka, India

Priya D S, JSS Academy of Technical Education, Bangalore, Karnataka, India

\*\*\*

**Abstract** - The examination of the Big Data phenomena is the focus of this essay. There are seven sections to it. In the first, it is explored how data and information are playing a larger and faster-growing role in the new socioeconomic reality. The concept of "Big Data" is then described, along with the primary factors driving data expansion. The most important options related to big data are given and addressed in the section that follows. The description of the tools, approaches, and the most beneficial data in the context of Big Data projects is the focus of the next section. The success aspects of Big Data efforts are examined in the section that follows, which is then followed by an examination of the most significant issues and difficulties related to big data. The paper's most important results and recommendations are presented in the final section.

**Key Words:** Big Data; Road block; Algorithms; Technology; Tools;

## 1. INTRODUCTION

Increasing amounts of data are entering modern businesses as a result of the volume of data being generated by those organizations' stakeholders and other entities operating in their business environments, in addition to the organizations themselves, which is expanding quickly. Thus, phrases like "a data-centric world" are becoming more and more common in this context.

The processes listed above are important components of the global socioeconomic shifts occurring now, where the extraordinarily dynamic growth of increasingly potent and widespread information technology plays a vital role. The modern economy has undergone tremendous change, and the establishment of a "interconnected economy" has been significantly accelerated by developments in this field. In terms of resources, this new type of economy is a knowledge-based one in which intellectual capital ranks as the most significant form of capital. Under these circumstances, the capacity of an organization to gather the appropriate data and information and to efficiently transform it into useful knowledge becomes an increasingly crucial issue.

Due to the acceleration of advancements made recently, the field of information technology has started to move into a new era. Processing power and data storage are almost free today, and networks and cloud-based services offer users

ubiquitous access to a wide range of services. These procedures lead to the creation of Big Data sets that have rapidly increased in size. Every day in 2012, 2.5 exabytes of data were produced, tripling from 2011. around every 40 months. In general, the previous two years have seen the creation of 90% of the world's data.

As a result, companies now have access to an unprecedented amount of data and information for analysis [4]. This creates a wide range of brand-new operational opportunities for firms as well as countless new difficulties.

The phrase "Big Data" has evolved in this context and is being used more frequently in the business world.

## 2. THE MOST SIGNIFICANT DATA GROWTH SOURCES AND BIG DATA

Because the term "Big Data" is not always understood and applied, other methods of analysis have been developed. The Leadership Council for Information Advantage claims that this phrase is not precise. "(...) it's a description of the never-ending assemblage of various types of data, the majority of which is unstructured. It speaks of data sets that are exponentially expanding and too big, unprocessed, or unstructured to be analysed with relational database approaches. Contrarily, Big Data is defined by NewVantage Partners as "a term used to describe data sets that are so large, so complex, or that require such rapid processing (sometimes called the Volume/Variety/Velocity problem), that they become challenging or impossible to work with using standard database management or analytical tools." It is crucial to emphasise to stress that Big Data refers to both the data associated with this consumption as well as the storage and consumption of original material.

In general, a number of important phenomena have led to a considerable increase in data generation.

The growth of traditional transactional databases, the first trend, is mostly related to the fact that businesses are gathering data more often and with greater granularity.

This is because of a number of factors, including rising client demands, escalating competition, and turbulent business environments. Organizations must respond to the changes occurring as quickly and flexible as possible, and then make necessary adjustments. To be able to accomplish this, they

are compelled to conduct ever-increasingly in-depth analyses of markets, rivalry, and customer behaviour.

The second trend, an increase in multimedia material, is related to the quick uptake of multimedia throughout a range of modern economic sectors, including the healthcare industry, where over 95% of clinical data is now digital video format has been generated. In general, multimedia data already makes up more than half of Internet backbone traffic, and by the end of 2013, it is anticipated that this share will reach 70%.

The advent of "The Internet of Things," a phenomenon where physical items or gadgets connect with one another without any human involvement, is the next trend that has contributed to an increase in the amount of data being generated. They connect to one another wirelessly or through wired connections, frequently utilising IP protocols. They gather and transmit enormous volumes of data because they are fitted with several sensors or actuators. The volume of data produced by the "Internet of Things" will increase dramatically by 2015, as the number of deployed connected nodes worldwide is anticipated to increase at a rate of over 30% each year.

The next very important source of the rise in data is social media. Just Facebook users produce a tonne of data. In 2011, Facebook's 600 million active users spent more than 9.3 billion hours per month on the site, producing an average of 90 pieces of content (pictures, notes, blog entries, links, or news articles). One billion people were using Facebook a year later. If only texts are taken into account, according to research done at the start of 2012, users send and receive an average of nine messages every month. In the instance of YouTube, 24 hours of video are uploaded per minute, while 98000 tweets are sent during the same period of time on Twitter. Smart phones are also becoming more and more significant in social networks. Social network usage is rising on both PCs and smartphones, but smartphone usage is substantially higher. If frequent users are taken into account, the percentage for PCs is 11% each year while the percentage for cellphones is 28% per year. Due to this, mobile data traffic has rapidly increased and doubled during the third quarters of 2011 and 2012. By 2018, mobile data traffic is expected to multiply twelve times.

### **3. OPPORTUNITIES AND BENEFITS RELATED TO THE USE OF BIG DATA**

The evolution of the Big Data phenomenon and the tools and procedures that go along with it cannot be divorced from the broader organisational changes that have been going on in recent years. In fact, it is becoming more and more prevalent in firms that are interested in the analytics sector and has greatly increased the possibilities offered by business intelligence (BI) technologies. Business intelligence systems are highly suited for gathering and analysing structured data

due to their role in offering firms a variety of possibilities and chances in the field of analytics. However, there are several analysis kinds that BI cannot handle. These mostly concern instances where data sets expanding in diversity, granularity, real-time, and iteration. When attempting to use conventional methodologies based on relational database models, these forms of unstructured, large volume, and rapidly changing data present challenges. It is now clear that a new class of technology and analytical techniques are in increasing demand.

The findings of research undertaken by the McKinsey Global Institute have verified that using Big Data has a variety of benefits depending on the sector of the economy. These findings demonstrate how Big Data has the power to transform industries as diverse as healthcare, public administration, retail, manufacturing, and personal location data. There are seven primary groupings of benefits associated with Big Data projects, according to the findings of a poll performed in the summer of 2012 by New Vantage Partners among C-level executives and department heads from many of America's leading firms. The most significant of these advantages are better, fact-based decision making (22%) and improved customer experiences (22%), along with the general message that the expectation is to take quicker, smarter decisions. Increased sales (15%), new product developments (11%), decreased risk (11%), more efficient operations (10%), and higher-quality goods and services (10%) are among the other groups of advantages.

Big Data platforms enable organisations to receive answers to critical queries instantly rather than over the course of months. Big Data's main benefit is to shorten the time it takes to get an answer, which speeds up decision-making at both the operational and tactical levels. The ability for ongoing business experimentation to inform decisions and test new goods, business models, and customer-focused innovations is a crucial new feature related to the Big Data phenomenon in the context of decision-making. Sometimes, a strategy like this even enables real-time decision-making. There are numerous instances of businesses adopting this in real life. For instance, Capital One's multidisciplinary teams do more than 65,000 tests annually. They play around with combining different market segments and fresh goods.

Based on online data streams, the online grocer FreshDirect changes prices and discounts every day or even more often. Tesco is another illustration. Through a loyalty card programme, this corporation collects transaction data on millions of its clients, which it then utilises to assess potential new business ventures. For instance, it examines how to tell customers about pricing, promotion, and shelf allocation decisions as well as how to design the most successful promotions for particular consumer segments. One such illustration is Walmart. The Big Data platform (The Online Marketing Platform) was developed by this company, and it is used, among other things, to conduct several

concurrent tests to evaluate new data models. Additionally, Internet powerhouses like Amazon, eBay, and Google have been leveraging to drive their performance through testing.

There are five main ways that big data adds value for organisations, according to the McKinsey Global Institute:

- enabling experimentation to identify needs, reveal variability, and enhance performance;
- fostering transparency by integrating data and making it more readily accessible to all pertinent parties;
- Segmenting populations to tailor interventions, Segmenting populations to tailor interventions,
- Using automated algorithms to supplement or replace human decision-making.
- Developing novel business models, goods, and services

Overall, the research findings from the Economist Intelligence Unit, which polled 607 executives from around the world in February 2012, support the importance of businesses utilising big data. According to the CEOs who participated in the poll, big data projects have raised their firms' performance over the past three years by about 26%. They anticipate that over the following three years, these initiatives will boost performance by an average of 41%. Additionally, it is important to note that businesses with decision-making processes based on data and business analytics have 5-6% higher output and productivity, according to the findings of research by Brynjolfsson et al. Business analytics-based decision-making also affects other performance metrics such as market value, equity return, and asset utilization.

Big Data systems have been utilised for both decision automation and human decision support, similar to BI initiatives. Based on the degree of risk associated with the decision, Big Data is utilised, on average, for decision support 58% of the time and for decision automation around 29% of the time, according to the findings of the aforementioned study done by the Economist Intelligence Unit.

#### **4. MAXIMUM USEFUL DATA, TOOLS, AND TECHNIQUES IN THE CONTEXT OF BIG DATA INITIATIVES**

The efficient implementation of big data initiatives necessitates adopting the proper organizational steps, such as ensuring that businesses have access to all the resources required to enable analysis of the always expanding data sets they have access to. One of the most important issues in this area is the application of appropriate methods and technologies. In reality, businesses combine, manipulate, analyse, and visualise big data using a wide range of techniques and technology. They come from a variety of

disciplines, including economics, computer science, statistics, and applied mathematics. Some of them have been specifically created for this, while others have been modified for it. Examples of methods used for Big Data analysis are: time series analysis, sentiment analysis, spatial analysis, simulation, data mining, data fusion and integration, A/B testing, and machine learning. Big Table, Cassandra, Google File System, Hadoop, Hbase, MapReduce, stream processing, and visualisation (tag cloud, clustergram, history flow, spatial information flow) are a few examples of the technologies used to gather, modify, manage, and analyse big data.

There are more and more brand-new analytical toolkits available for the examination of Big Data. Examples of these remedies include:

- NM Incite, Social Mention, SocMetrics, Traackr, Tweepi (sentiment analysis tools for estimating the buzz around a product or service, influencer intelligence tools for identifying key influencers and targeting for marketing or insights)
- Attensity, Autonomy (live testing tools for getting direct user feedback on new products)
- Alterian, TweetReach (network intelligence tools for real-time analysis of the reactions and responses to changes of industry players),

In addition, adequately trained personnel are a crucial component of Big Data initiatives. In this context, a special category of worker known as a data scientist who has received the necessary training to deal with Big Data is mentioned. In actuality, it means that they should be adept at mining vast amounts of unstructured data for the solutions to an organization's most pressing problems. These individuals ought to combine the skills of an analyst, a data hacker, a communicator, and a trusted advisor. They should have strong analytical capabilities as well as strong, innovative IT skills, and they should be familiar with the company's internal products and operations. The majority of businesses use platforms to bridge the knowledge gap because data scientists often require years to acquire in-depth domain expertise.

Appropriate data is a fundamental resource needed for Big Data initiatives, in addition to appropriate methodology, tools, and people. As was already noted, modern enterprises are currently receiving a large amount of data from numerous sources, but not all Big Data sets are equally valuable. The most crucial source of data is unquestionably information about business activity, such as sales, purchases, costs, and so forth. The second major source of data is office documents, closely followed by social media. In several industries, like healthcare, pharmaceuticals, and biotechnology, social media data sets are more significant than office documentation. POS data, website clickstream



data, RFID/logistics data, GIS data, telecommunications data, and telemetry data are some of the additional crucial types of data sets.

## 5. REASONS BIG DATA INITIATIVES SUCCESS

Numerous success elements can be identified through an examination of Big Data efforts that have already been put into action, each with its own set of suggestions.

Five crucial rules have been highlighted by Marchand and Peppard as essential to a Big Data project's success. They consist of:

1. Making the Big Data initiative's focus on people.
2. Putting a focus on information use as a means of maximising the value of information technology.
3. Including cognitive and behavioural scientists on IT project teams.
4. Concentrating on education
5. Focusing on using technology rather than resolving business issues.

Barton and Court, on the other hand, came to the opinion that complete exploitation of data and analytics needs three competences based on their experiences working with businesses in data-rich industries:

1. Picking the appropriate data. In this environment, it's crucial to upgrade IT architecture and infrastructure for simple data merging as well as source internal and external data creatively.
2. Stressing the use of information as a key to maximising the potential of information technology. In this setting, it's crucial to concentrate on the performance factors that matter most and to create models that strike a balance between complexity and usability.
3. Including cognitive and behavioural scientists on IT project teams. In this context, upgrading business processes and developing capacities to enable tool use are two crucial factors. The first is building straightforward, intelligible tools for workers on the front lines.

Organizations that utilise Big Data base their operations on three key concerns, according to Barth et al. are:

1. Focusing on data flow rather than stock prices
2. Relying on product and process developers and data scientists rather than data analysts.

3. Bringing operational and production processes into the centre of the business and separating analytics from the IT function.

## 6. THE BIG DATA REFERENCES' MOST SIGNIFICANT OBSTACLES AND CHALLENGES

Big Data initiatives, like other IT-related endeavours, are not without issues and difficulties. The research by the Economic Intelligence Unit referenced previously identifies some of the challenges to effectively using big data for decision-making. The biggest hurdle was "organizational silos" (55,7%), which are caused by the fact that data related to specific organizational functions (such as sales, distribution, etc.) are collected in "function silos" rather than being pooled for the benefit of the entire business. The second problem, which is also significant (50,6%), is the dearth of data scientists with the necessary qualifications. The third factor (43,7%) is how long it takes businesses to analyse large data sets. The third factor (43,7%) is how long it takes businesses to analyse large data sets. Organizations anticipate being able to examine data in real time and take action on it, as was already said. The challenges associated with analysing ever-increasing amounts of unstructured data make up the fourth barrier (41,7%). The fifth major barrier is, finally, senior management's failure to view big data in a sufficiently strategic way (34,9%).

Five management issues, according to McAfee and Brynjolfsson, are holding back firms from utilising big data to its full potential. They are: company culture, leadership, people management, technology, and decision-making. Having more or better data does not ensure success when evaluating leadership. The firm's management still need to have a clear understanding of the market, set attainable goals, and see how the organization will develop. Numerous organizational decisions are altered by big data. The need to provide the organization with the necessary individuals (such as data scientists) who are capable of working with large amounts of data is related to talent management. The next issue is how to guarantee that the data scientists have the right equipment to handle Big Data. Technology is a crucial component of big data endeavours, even though it is not sufficient on its own to be successful. The second obstacle is related to the issue of making sure there exist mechanisms in place to ensure that the information and the appropriate decision-makers are present in the same place. Making ensuring that those who comprehend the issues can make the appropriate use of information and collaborate with those who possess the essential problem-solving abilities is crucial. The last difficulty is linked to alterations in organizational culture. Making decisions that are as data-driven as possible rather than relying solely on intuition and gut feeling is crucial in this situation. It is important to note that other research, such as that pertaining to sectoral Big Data projects, has also discussed the significance of such cultural transformation.

It's also important to consider the numerous difficulties related to data and legal rights. They have to do with things like copyright, database rights, privacy, trade marks, contract law, and competition law. Another significant issue relating to legal matters also exists. It has to do with the openness of data collection procedures. The use of Big Data to boost decision automation is a further significant risk. One further significant risk is highlighted in the context of big data. It has to do with the possibility that Big Data isn't giving a certain circumstance the full picture. This is due to a number of factors, including biases in the data gathering process, exclusions or gaps in the data signals, or the ongoing requirement for context in conclusions.

At the same time, pre-Big Data dangers and difficulties are still evolving, such as the issue of protecting information and data that has been gathered. These problems primarily concern how to safeguard information that must be kept private by businesses (such as different categories of consumer data) and information that is competitively sensitive. As a result, the issues related to broadly defined security of enterprises' IT infrastructure and protection against various assaults become even more crucial than before. Securing it has become even more crucial as a result of enterprises' growing reliance on the efficient and stable operation of their IT infrastructure as a result of the Big Data phenomena.

## 7. CONCLUSIONS

The processes pertaining to decision-making at different organizational levels are evolving as a result of the fast expanding amount of data that companies have at their disposal and the opportunities connected with its practical usage. Big Data thus has the ability to significantly improve how businesses operate generally and gives them a competitive edge. Now, businesses are attempting to make even greater use of the possibilities and opportunities that are appearing.

It is not sufficient to simply collect and own the necessary data sets if projects aimed at the actual exploitation of Big Data sets are to be successful at providing an organization with a competitive edge and be of value. In actuality, this is just where every Big Data endeavor begins. Additional crucial components are appropriate analytical models, tools, qualified personnel, and organizational skills. When any one of these essential elements is absent, the result could be that instead of the anticipated benefits, there is merely disappointment. Big Data projects are merely the latest in a long history of managerial fads, leading to disillusionment and the notion. In general, the speed at which Big Data solutions can be used in a secure and practical way remains questionable, despite the fact that they offer enormous potential for both commercial enterprises and governments.

## REFERENCES

- [1] Kemp Little LLP, "Big Data - Legal Rights and Obligations", <http://www.kemplittle.com/Publications/WhitePapers/Big%20Data%20Legal%20Rights%20and%20Obligations%202013.pdf>, January 2013.
- [2] D. Tapscot, A. Williams, *Radical Openness*. New York: TED Books (Kindle Edition), 2013.
- [3] National Intelligence Council, "Global Trends 2030", <http://globaltrends2030.files.wordpress.com/2012/11/global-trends-2030-november2012.pdf>, December 2012.
- [4] E. Brynjolfsson, A. McAfee, (2012), "Big data: The management revolution", *Harvard Business Review*, pp. 60-68, October 2012.
- [5] A. Lampitt, "Hadoop: Analysis at massive scale", in *InfoWorld*, [http://resources.idgenterprise.com/original/AST-0084522\\_IW\\_BigData\\_rerun\\_1\\_all\\_sm.pdf](http://resources.idgenterprise.com/original/AST-0084522_IW_BigData_rerun_1_all_sm.pdf), pp. 8-12, Winter 2013.
- [6] LCIA, "Big Data: Big Opportunities to Create Business Value", <http://poland.emc.com/microsites/cio/articles/big-data-big-opportunities> / LCIA-BigData-Opportunities-Value.pdf, 2011.
- [7] McKinsey Global Institute, "Big data: The next frontier for innovation, competition, and productivity", [http://www.mckinsey.com/mgi/publications/big\\_data/pdfs/MGI\\_big\\_data\\_full\\_report.pdf](http://www.mckinsey.com/mgi/publications/big_data/pdfs/MGI_big_data_full_report.pdf), May 2011.
- [8] NewVantage Partners, "Big Data Executive Survey: Themes & Trends", <http://newvantage.com/wp-content/uploads/2012/12/NVP-Big-Data-Survey-Themes-Trends.pdf>, 2012.
- [9] J. Gantz, D. Reinsel, "Extracting Value from Chaos"