

Big Data Analytics: A Concept

Ms. Pooja Dadhe¹ and Dr. R.N. Jugele²

¹(poojadadhe@gmail.com, Department of Computer Science, Shivaji Science College, Nagpur, India)

²(rn_jugele@yahoo.com, Department of Computer Science, Shivaji Science College, Nagpur, India)

Abstract: Big data is a term for massive data sets having large, more varied and complex structure with the difficulties of storing, analyzing and visualizing for further processes or results. Big data is not only changing the way we store data. The process of research into massive amounts of data to reveal hidden patterns and secret correlations named as big data analytics. These useful information's for companies or organizations with the help of gaining richer and deeper insights and getting an advantage over the competition. Big data implementations need to be analyzed and executed as accurately as possible. This paper presents an overview of big data's content, importance, work and key technologies.

Keywords – Big Data, Big data Analytics, information, Hadoop, Predictive Analytics, Text Mining.

I. INTRODUCTION

Big data is in every industry and business function and is an important factor for production. MGI estimated that 7 Exabyte's of data enterprises globally were stored in 2010. Interestingly, more than 50 % of IP traffic is non human and M2M will become increasingly important [1]. So what is big data suppose to create? Value. But "value" exactly? David Gorbet of Mark Logic [2] explains: "the increase in data complexity is the biggest challenge that every IT department and CIO must address. Business across industries stores the data also able to leverage it quickly and effectively to derive business value. Value comes only from what we infer from it and why we need big data analytics.

II. WHAT IS BIG DATA ANALYTICS?

Big data analytics examines large amounts of data to uncover hidden patterns, correlations and other insights. With today's technology, it's possible to analyze data and get answers from it immediately – an effort that's slower and less efficient with more traditional business intelligence solutions.

III. HISTORY AND EVOLUTION OF BIG DATA ANALYTICS

The concept of big data has been around for years; most organizations now understand that if they capture all the data that streams into their businesses, they can apply analytics and get significant value from it. But even in the 1950s, decades before anyone uttered the term "big data," businesses were using basic analytics to uncover insights and trends.

The new benefits that big data analytics brings to the table, however, are speed and efficiency. Whereas a few years ago a business would have gathered information, run analytics and unearthed information that could be used for future decisions, today that business can identify insights for immediate decisions. The ability to work faster – and stay agile – gives organizations a competitive edge they didn't have before.

IV. WHY IS BIG DATA ANALYTICS IMPORTANT?

Big data analytics helps organizations harness their data and use it to identify new opportunities leads to smarter business moves, more efficient operations, higher profits and happier customers. In his report *Big Data in Big Companies*, IIA Director of Research Tom Davenport interviewed more than 50 businesses to understand how they used big data. He found they got value in the following ways:

1. **Cost reduction.** Big data technologies such as Hadoop and cloud-based analytics bring significant cost advantages when it comes to storing large amounts of data and plus they can identify more efficient ways of doing business.
2. **Faster, better decision making.** With the speed of Hadoop and in-memory analytics, combined with the ability to analyze new sources of data, businesses are able to analyze information immediately which makes decisions based on what they've learned.

3. **New products and services.** With the ability to gauge customer needs and satisfaction through analytics comes the power to give customers what they want. Davenport points out that with big data analytics, more companies are creating new products to meet customers' needs.

V. BIG DATA ANALYTICS IN TODAY'S WORLD

Most organizations have big data. And many understand the need to harness that data and extract value from it. These resources cover the latest thinking on the intersection of big data and analytics. High-performance analytics lets user do things you never thought about before because the data volumes were just way too big. For instance, it can get timely insights to make decisions about fleeting opportunities, get precise answers for hard-to-solve problems and uncover new growth opportunities and using while using IT resources more effectively

Who is using it?

1.1. Travel and hospitality.

Keeping customers happy is key to the travel and hotel industry, but customer satisfaction can be hard to gauge in a timely manner. Resorts and casinos have only a short window of opportunity to turn around a customer experience that's going south fast. Big data analytics gives these businesses the ability to collect customer data, apply analytics and immediately identify potential problems before it's too late.

1.2. Health care

Big data is a given in the health care industry. Patient records, health plans, insurance information whereas other types of information is difficult to manage – but are full of key insights once analytics are applied. That's why big data analytics technology is so important to health care. By analyzing large amounts of information – both structured and unstructured, health care providers provide lifesaving diagnoses or treatment options almost immediately.

1.3. Government

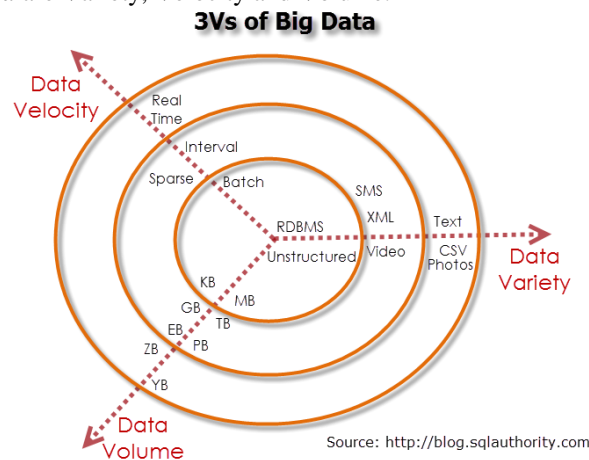
Certain government agencies face a big challenge: tighten the budget without compromising quality or productivity. This is particularly troublesome with law enforcement agencies, which are struggling to keep crime rates down with relatively scarce resources. Many agencies use big data analytics; the technology streamlines operations while giving the agency a more holistic view of criminal activity.

1.4. Retail

Customer service has evolved in the past several years, as savvy shoppers expect retailers to understand exactly what they need, when they need it. Big data analytics technology helps retailers meet those demands. Armed with endless amounts of data from customer loyalty programs, buying habits and other sources, retailers have an in-depth understanding of their customers, they can also predict trends, recommend new products and boost profitability.

VI. THE 3 V'S OF BIG DATA

The 3Vs that define *Big Data* are **Variety**, **Velocity** and **Volume**.



2.1 Volume

The most visible aspect of Big Data, referring to the fact that the amount of generated data has increased tremendously the past years. This is the less challenging aspect in practice.

2.2 Velocity

The data growth and social media explosion have changed how we look at the data. There was a time when user used to believe that data of yesterday is recent. The matter of the fact newspapers is still following that logic. However, news channels and radios have changed how fast receive the news. Today, people reply on social media to update them with the latest happening. On social media sometimes a few seconds old messages is not something interests users. They often discard old messages and pay attention to recent updates. The data movement is now almost real time and the update window has reduced to fractions of the seconds. This high velocity data represent **Big Data**.

2.3 Variety

Data is stored in multiple formats. database, excel, csv, access or for the matter of the fact and stored in a simple text file. Sometimes the data is not even in the traditional format it may be in the form of video, SMS, pdf. It is the need of the organization to arrange it and make it meaningful. It will be easy to do so if we have data in the same format, however it is not the case most of the time. The real world has data in many different formats and that is the challenge need to overcome with the **Big Data**. This variety of the data represent represents **Big Data**.

VI. HOW IT WORKS AND KEY TECHNOLOGIES

There's no single technology that encompasses big data analytics. There's advanced analytics that are applied to big data, but in reality several types of technology work together to help you get the most value from your information. Here are the biggest players:

Data management: Data needs to be high quality and well-governed before it reliably analyzed. With data constantly flowing in and out of an organization, it's important to establish repeatable processes to build and maintain standards for data quality. Once data is reliable, organizations establish a master data management program that gets the entire enterprise on the same page.

Data mining: Data mining technology helps to examine large amounts of data to discover patterns in the data – and this information used for further analysis to help answer complex business questions. With data mining software can sift through all the chaotic and repetitive noise in data, pinpoint what's relevant, use that information to assess likely outcomes, and then accelerate the pace of making informed decisions.

Hadoop: This open source software framework store large amounts of data and run applications on clusters of commodity hardware. It has become a key technology to doing business due to the constant increase of data volumes varieties, and its distributed computing model processes big data fast. An additional benefit is that Hadoop open source framework is free and uses commodity hardware to store large quantities of data.

In-memory analytics: By analyzing data from system memory, it derive immediate insights from your data and act on them quickly. This technology is able to remove data prep and analytical processing latencies to test new scenarios and create models; it's not only an easy way for organizations to stay agile and make better business decisions, it also enables to run iterative and interactive analytics scenarios.

Predictive analytics. Predictive analytics technology uses data, statistical algorithms and machine-learning techniques to identify the likelihood of future outcomes based on historical data. It's all about providing a best assessment on what will happen in the future, so organizations feel more confident that they're making the best possible business decision. Some of the most common applications of predictive analytics include fraud detection, risk, operations and marketing.

Text mining: Text mining technology analyzes text data from the web, comment fields, books and other text-based sources to uncover insights. Text mining uses machine learning or natural language processing technology to comb through documents – emails, blogs, Twitter feeds, surveys, competitive intelligence and help to analyze large amounts of information and discover new topics and term relationships.

VII. CONCLUSION:

Big Data is not just about lots of data, it is actually a concept providing an opportunity to find new insight into existing data guidelines to capture and analysis your future data. It makes any business more agile robust to overcome business challenges. Big data poses opportunities and challenges for businesses. Previously untapped sources of data are able to be stored and processed. Benefits of big data use to business executives include enhanced data sharing through transparency, improved performance though analysis, augmented market segmentation, increased decision support through advanced analytics, and greater ability to innovate products, services and business models. Business owners need to follow trends in big data carefully to make the decision that fits their businesses.

REFERENCES

Journal Papers:

- [1]. Neerja Kulkarni, Dr. Priti Puri, Big Data Analytics – Business Benefits and Technical Challenges, *International Journal of Modern Computer Science and Applications (IJMCSA)*, Volume No.-3, Issue No.-2, April, 201, ISSN: 2321-2632 (Online)

Books:

- [2]. Rajendra Akerkar, *big data computing*, (CRC Press,NewYork, 2014)
[3]. Paul Zikopolous, Dirk DeRoos, Chirs Eaton, Tom Deutsch, George Lapis, Understanding big data,Analytics for Enterprise class Hadoop and streming class (McGraw hill Publications,Newyork 2012)

Chapters in Books:

- [4]. James Manyika, Michael Chui, McKinsey Global Institute, Big data: The next frontier for innovation, competition, and productivity, May 2011