

# Big Data Analytics – From Data to Insights

Jaya Singh, Prof (Dr) Ajay Rana

**Abstract**— The expanding digital world of Web 2.0, sensors, digital communications and Cloud services has created a spate of data globally which is continually growing in volume at a streaming velocity and exist in structured and unstructured formats. This data is very well-known worldwide as Big Data and has become a source of ample opportunities for the enterprises. Business analysts, data experts and surveys carried out across organizations have proved that the power of big data is amazing and it can help the enterprises to innovate and gain competitive edge by extracting the insights from big data and applying it for decision-making and analysing the customers. Advanced analytics is used to glean information from the big data as traditional business intelligence techniques may not be sufficient to deal with the unstructured content of big data. However, adoption of big data analytics is not an easy task and calls for a sound management strategy. This paper discusses the evolution of big data having distinguishing characteristics and discusses the surge and importance of big data analytics and the approach to embrace the revolution of analytics in the enterprises.

**Index Terms**— big data, advanced analytics, enterprise, business intelligence, data-driven, insights

## 1 INTRODUCTION

**D**ATA pouring from social media like Facebook, Twitter, LinkedIn, YouTube, clickstreams, web logs, emails, documents, videos and audios from mobile phone users and cloud applications is known as big data. The data is termed as “big” as it is huge in volume, which is continually growing at an astonishing rate due to the falling cost of computing and the exposure of an increasing number of users to the digital technology. This data does not have a fixed schema due to its unstructured format and hence traditional relational databases cannot be used to capture and store it. That’s how Hadoop has become one of the popular choices in dealing with big data. Hadoop is an open –source technology and stores data in various distributed clusters. Some other technologies are NewSQL, NoSQL, in-memory databases and to be precise big data technology is still evolving [1].

Top researchers in business and academia have predicted that big data needs to be mined by applying advanced analytics consisting of sophisticated data mining algorithms, statistical analysis and complex mathematical calculations. The insights hence gained must be then applied in decision-making and innovations. Big data requires an iterative approach on raw data to achieve high quality understanding [2]. Big Data analytics requires tools and technologies to collect, ingest and analyse the big data and the skills to work with these emerging technologies. So, the right techniques with the right talent at the right time are required to perform analytics. There should be an iterative exploration and investigation of raw data and the insights gained must be applied on business transformation and making effective decisions to gain a competitive edge in the market.

### 1.1 The Era of Big Data

Big Data has made its presence strongly felt in the enterprises and economy and the society. There are two types of data present in the enterprises - one, transactional data from ERP, CRM, SCM stored in relational data bases and second, the unstructured data from disruptive technologies including Web 2.0, sensors, social media and cloud services [3]. Emails, documents, videos, audios and texts are being generated at tremendous fast rate. Facebook has 500 million users and this

number is on increase day-by-day and these users are uploading videos, writing blogs and use many other features leading to an enormous increase in the volume of data. And, since the unstructured data like videos and audios don’t have fixed schema, normal relational data bases having stable schema, are not competent enough to handle it [4]. Column-oriented databases, in-memory databases and multi-scalable distributive technologies are required to tackle big data.

This needs the enterprises to invest in the necessary infrastructure required to ingest and analyse big data and the major concern here is the integration of legacy data with the big data [5].

### 1.2 What Makes Big Data Different?

Several dimensions associated with the big data make it different from legacy data. These aspects are tabulated in the table 1.

**Table 1. Dimensions of Big Data [1]**

S. No.	Dimension	Description
1.	Volume	Big data is continuously generated and the rate of data expansion is estimated to be 40% per annum due to emergence of sensors, social media and mobile computing [1, 3].
2.	Velocity	Data is being created at lightning speed due to sophistication in the digital technologies.
3.	Variety	Big data has diverse formats due to its unstructured content and has unstable schema.
4.	Value	Big Data brings along with it a plethora of opportunities which need to be identified and leveraged in the enterprises for innovation, productivity and competence.
5.	Complexity	The volume, velocity and variety of big data make it complex to handle.

## 2 SURGE OF BIG DATA ANALYTICS

The big data has arrived in each and every sphere of the economy. Researchers have started realizing that this data landscape must be understood using advanced technologies and the information gained must be used wisely for business transformation. This is how the big data analytics surged up. Businesses had been analysing the data present with them to improve the market position and understand the market. However, the data in the era of big data is different from the traditional structured data and requires a new approach [6]. The analysts are following an iterative approach to glean better and better information by creating more meaningful queries to the results produced in every cycle.

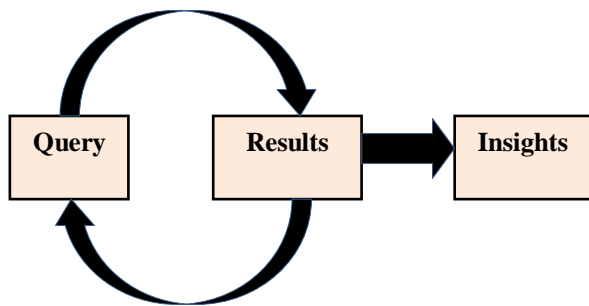


Figure 1. An Iterative approach to Analytics [3]

The analysts are following an iterative approach to glean better and better information by creating more meaningful queries to the results produced in every cycle. It has also been predicted by the researchers and business experts that the global economy and the society will be driven by an inevitable rise of big data analytics in the coming years. The big data market's growth can be depicted using the figure below.

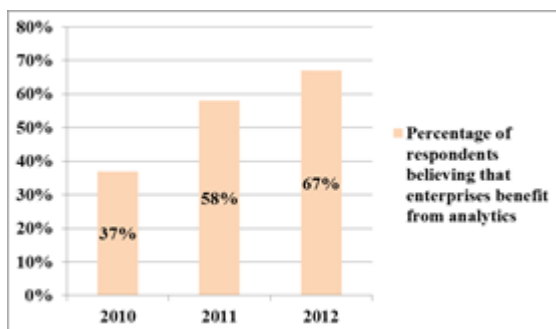


Figure 2. Growing Competitive With Analytics [2]

### 2.1 Defining Big Data Analytics

Big Data analytics refers to the skills, technologies, applications and practices for a continuous and iterative exploration and investigation of raw data and extract meaningful information and put these insights into action for better decision-making and transforming the businesses [7, 8]. Enterprises practising advanced analytics like predictive and indicative modelling have proved that success worth billion dollars is associated with the implementation of big data analytics solu-

tion. However, the path to big data analytics is not an easy task as the enterprises must have a sound management strategy to take care of the concerns associated with it.

Big Data analytics offers several benefits like [3, 6]:

- Extraction of hidden patterns and trends within the data landscape
- Identification for new opportunities leading to innovation
- Creation of business scenarios
- Making predictions about future business
- Prescription of actions

### 2.2 Requirements of Big Data Analytics

Big Data Analytics applies advanced techniques like statistical analysis, predictive modelling, sentiment analysis, text analysis, complex queries using new technologies like NewSQL, NoSQL and deep mathematical calculations. Hadoop has emerged as one of the popular storage techniques to ingest and analyse the big data. However, Hadoop is not the only one solution for big data. Traditional data from ERP, CRM, SCM stored in relational databases must be integrated with the big data and this has been seen as the biggest hurdle by the enterprises. Since these and other evolving technologies require highly talented professionals to work with them, getting right skills is also a major concern [9].

## 3 ADOPTION OF BIG DATA ANALYTICS IN THE ENTERPRISE

The enterprise IT has always been striving to extract the pearls of information from the ocean of data quickly. The current scenario in the world of data is different from what existed earlier. Data now is created not only from the legacy sources present within the premises of the enterprise such as ERP, CRM, SCM, emails but also from external sources such as sensors, social media, mobile computing and cloud services. And analysis of unstructured content like videos, audios, sensor data etc. is a tough task as these don't have a fixed structure.

Analysis of internal as well as external data in the organizations did not exist earlier due to limited technologies and awareness. With the fall in cost of computing and evolution of new technologies like Hadoop, MapReduce, in-memory databases, NoSQL, NewSQL, it is already in use.

However, the enterprises are embracing the big data analytics and many of them have been extremely beneficial. These enterprises hail from the fields of health care, education, transportation and finance. These enterprises have followed a sound strategy for data governance, quality and security issues [10, 12].

### 3.1 Enterprise Approach to Big Data Analytics

Enterprises must deal with the big data by cultivating a data-driven culture among the employees. Organizations must not operate in siloes creating information islands and must consider data as a shared asset. Also, a sound management strategy must be established within the enterprise to tackle big data flooding from in-house applications and cloud services

and several third party sources.

According to a survey conducted by MIT Sloan Management Review team, enterprises implementing analytics are continually innovating [2, 11]. The organizational culture plays a vital role in deciding whether the organization will be able to achieve competitive advantage or not, besides the tools and skills required to implement analytics. Most advanced users of analytics have a robust data-oriented culture flourishing within the enterprise [13]. The much talked about data-driven culture is nothing but is a pattern of behaviours and practices in the enterprise where data is considered as a shared resource and valued as much as other assets of the organization. Also, there exists a faith that business can be more innovative, productive and competitive with the insights gleaned from the data.

A data-centric enterprise can look beyond the pile of big data and gain meaningful outcomes and the data improves in a data-centric enterprise. However, this doesn't come so easily as it requires the way people in the enterprise used to perform their duties, had a thought process which needs modification. That is, human thinking needs to be changed by proper counselling by the senior management and it's not be forgotten that it is psychological for any human to resist change. The employee resistance had always been considered to be a very sensitive issue in the success of any organization and needs to be handled in a very sophisticated manner.

### 3.2 Step-by-step technique to implement big data analytics in the enterprise

The enterprise can approach the implementation of big data analytics to tap the power of unstructured data. Here is the step-by-step way an enterprise can move towards the world of big data analytics [12].

**Step 1:** Identification of core business requirements and creation of an enterprise-wide big data blueprint.

**Step 2:** Customization of infrastructure to meet the requirements of big data analytics.

**Step 3:** Data sources prevailing in the enterprise must be defined and their location, formats, way of tracking and replication technique should be identified.

**Step 4:** Defining known use cases, including the analytical models required to understand the aspects of data.

**Step 5:** Building analytic capabilities based on business priorities.

**Step 6:** Consideration of frequent, flouted issues of data performance, security and governance necessary for an efficacious implementation.

**Step 7:** Evaluation of the analytics tools and technology like predictive modelling, analytical modelling, text analysis, sentiment analysis, video analysis, in terms of features and functions.

**Step 8:** Definition of metrics of success necessary to evaluate the Return-On-Investment on the opted technology.

**Step 9:** Creation of a road map for the proposal mentioning the way big data analytics will be leveraged in the current business scenario and in the future.

Implementing advanced analytics is the only way to unveil the trends, patterns and opportunities hidden within the pile

of data. It is well-said that if big data is the challenge, big data analytics is the answer [14].

## 4 CONCLUSION

Evolving internet technologies accessible easily to users, lowering of computing cost and pacing digital communications have forced the enterprises to embrace the unfathomable data deluge and exploit it to the fullest with the help of emerging technologies like Hadoop, MapReduce, NoSQL etc. The age of big data has begun in its full swing. The paper described the multi-facet big data due to its different characteristics of volume, velocity and variety. The rise of data analysis and the role of data scientists is cropping up to get along with the big data analytics and converting the heap of data into actions for business transformation. The paper defined big data analytics and discussed how the IT world is revolutionized with the presence of analytics and motivated to implement it. The message is very clear – enterprises have big data, big data is a valuable asset with plethora of hidden opportunities, analytics is used to extract these pearls from big data and the information so obtained must be applied on the day-to-day operations of the business, yielding amazing results. As the big data technology stack is new and emerging, the organizations must move towards a data-centric approach, which is discussed in detail by the authors in the preceding sections. The paper also specifies a systematic approach to implement an effective analytics solution in the enterprise.

## REFERENCES

- [1] Eric Sweden, "Is Big Data a Big Deal for State Governments? The Big Data Revolution – Impacts for State Government – Timing is Everything", NAS-CIO Series on Analytics, August 2012.
- [2] David Kiron, Renee Boucher Ferguson, Pamela Kirk Prentice, "From Value to Vision: Reimagining the Possible with Data Analytics", MIT Sloan Management Review Research Report, Spring 2013.
- [3] HP Business White paper, "The disruptive power of Big Data – How Big Data analytics can transform your business", available online at <http://www.hp.com/>.
- [4] Jaya Singh, Ajay Rana, "Exploring the Big Data Spectrum", IJETAE, Volume 3, Issue 4, April 2013, pp. 73-76.
- [5] Dave Ohara, "The big machine: creating value out of machine-driven big data", GigaOM PRO, April 2012.
- [6] David S. Linthicum, "Big Data Analytics Deep Dive Deriving Meaning From the Data Explosion", InfoWorld Special Report sponsored by IBM.
- [7] David F. Carr, Joab Jackson, Thor Olavsrud, Bob Violino, "Strategic Guide to Big Data Analytics", A Strategic Guide from the editors of CIO.
- [8] IDC Analyze the Future, "Here Comes Big Data: Perspectives from Indian Enterprises", 2012, Commissioned by NetApp to IDC.
- [9] Michael Shroeck, Rebecca Shockley, Dr. Janet Smart, Professor Dolores Romero, Professor Peter Tufano, "Analytics: The real-world use of big data – How innovative enterprises extract value from uncertain data", IBM Institute for Business Value Executive Report, October 2012.
- [10] George Lumpkin, "Integrate for Insight", Oracle White Paper
- [11] Intel IT Center, "Big Data 101: Unstructured Data Analytics, A Crash Course on the IT Landscape for Big Data and Emerging Technologies", June 2012.
- [12] Interview with Tim McGuire, Mathew Ariker, Matthias Roggendorf, "Making Data Analytics Work: Three Key Challenges", McKinsey & Company,

March 2013.

- [13] Hsinchun Chen, Roger H. L. Chiang, Veda C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly, Vol. 36, No. 4/ December 2012.
- [14] Michael Hagstrom, Jim Davis, Cynthia Beath, Tony Fisher, "Big analytics brings big opportunities", Intelligence Quarterly Journal of Advanced Analytics IQ 2013

IJSER