



## **An Emerging Trend of Big data for High Volume and Varieties of Data to Search of Agricultural Data**

**PARAG SHUKLA, BANKIM RADADIYA and KISHOR ATKOTIYA**

<sup>1</sup>Head, Department of MCA, Atmiya Institute of Technology & Science, Rajkot – 360005, India.

<sup>2</sup>IT Director, Navsari Agricultural University, Navsari - 396445, India.

<sup>3</sup>Head, J. H. Bhalodiya Women's College, Rajkot- 360005, India.

(Received: June 10, 2015; Accepted: August 02, 2015)

### **ABSTRACT**

The data in amount of large size in world is growing day by day. Data is growing because of use of digitization, internet, smart cell and social media and networking. A collection of data sets which is very large in size as well as complex is called Big Data which is use in current trend and next generation data transformation, data analysis and data storage of agricultural crops, seeds, works, labors, tools and environment data. Generally before current scenario, size of the data is measure in Megabyte and Gigabyte but now a day it measure in Petabyte and Exabyte. Traditional database systems are not able to capture, store and analyze this large scale of data. As the internet is growing, amount of big data continue to grow. Big data analytics provide new ways for businesses updates and requirement for updating and government to analyze unstructured data. Now a days, Big data is one of the most important and challenging point in information technology world. It is executing very important role in future. Big data changes the way of world for management and use big amount of data. Some of the applications are in areas such as medical issues, healthcare, traffic issues, banking management, retail management, education and so on. Organizations are becoming more reliable & flexible and more open. Figure 2 is display a big data and analytics road map for large amount of data analysis and storage.

**Key words:** Big data, Faster analyzing of high volume data, Verities of data, Next generation, Big amount data search, No Sql database, Traditional database, Big data analysis, Volume, Verities of data, Velocity, Veracity, Agricultural environment.

### **INTRODUCTION**

In the past, the types of information available were limited. Approach to technology has a well-defined set Information management. But in today's world, our world has been the explosion of data volumes. It is Terabytes and petabytes. The large existing data from different databases

relational data stored in Big Data. It means a collection of large data sets is very complicated, they are. Big data can be described by the following terms:

#### **Volume**

Some small organizations may have gigabytes or terabytes of data storage. Regardless

of the amount of data would be the size of the organization continues to grow. Much of the data of these companies establish petabytes or even exabytes soon as they could get, but are now within reach terabytes. Volume data generated by the machine is greater than the conventional data.

### Verities

Different types of data are captured. It can be structured, unstructured or semi structures. Management is important to analyze further data types and file refers to many different, but for which traditional relational data bases are unsuitable. Examples of this type of sound and movie files, images, documents, geo-location data, weblogs, strings, web content, etc<sup>1</sup>.

### Velocity

As a flows of data from the data arriving continuously. The rate of change in the data and the speed that should be used to create real value is<sup>2</sup>.

### Veracity

Having large amounts of data is not correct; you can create a problem and cannot be of any use. Therefore, it must be right.

There are three types of large data - structured, unstructured and semi-structured. Similar institutions are grouped into structured data. The entities of a group have the same description. Examples of relational databases and spreadsheets are examples of information data structure data.

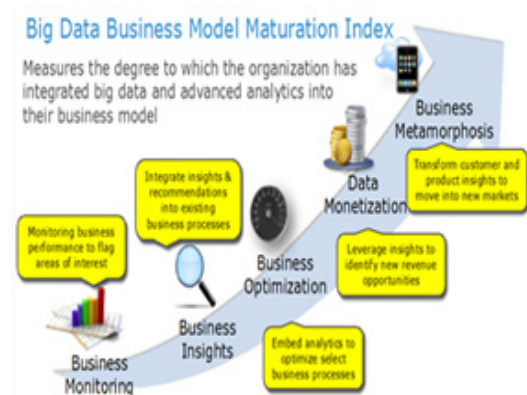


Fig. 1: Is shown a big data business model maturation index for large amount of data analysis in future of business

Unstructured whole numbers, words, figures, etc. Data can be of any kind and does not follow any rules. You cannot analyze with standard statistical methods. For large data, different tools are required. For example, social networks, e-mail, photographs, multimedia, etc. They are semi-structured data, similar entities are grouped together. The entities of the same group cannot have the same characteristic. Emails, EDI data are examples of this type.

Big Data is receiving much attention for the day. This new growth opportunities and help create new categories of business completely. The production process is combined with intelligence. Continuously improve processing power and new technologies for data analysis "Big data" that can be made from a variety of media sources. Creating great information source data so obvious or did not intend to make information about the data of all organizations. Big data is used properly, it can be a good idea to venture into business. Some of the applications of the systems are for emerging agricultural crop care, health care and many more.

Large data can come from different sources. It may be digitally generated and stored using a series of ones and zeros that can be, and can be manipulated by computer. This period or return of our daily life location data of mobile phones or digital service time may or negotiate. The data generated by social networking sites such as data

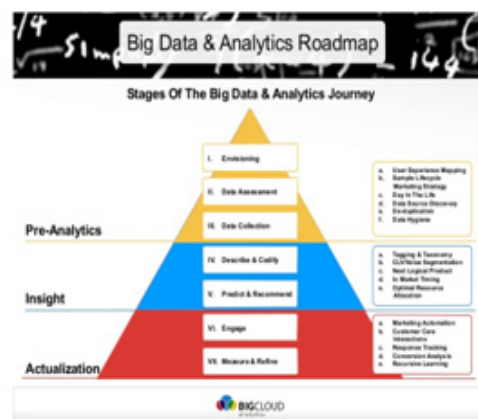


Fig. 2: Is display a big data and analytics road map for large amount of data analysis and storage

entry, RFID, sensor networks, etc. They can be generated using methods unorthodox out.

### Difference between traditional and big data analytics

Analysis of large data architectures is in may differ from traditional data sources traditional data processing that are structured internally. Data integration tools are used to extract, transform and load data from transactional databases. Then the quality of data and data normalization and data in rows and columns modeled. Modeling data is loaded into an enterprise data warehouse. Big Data using traditional methods to process data is too large. As the volume of data explodes, organizations such as social media data from different sources, such as the treatment of large data is not automatically

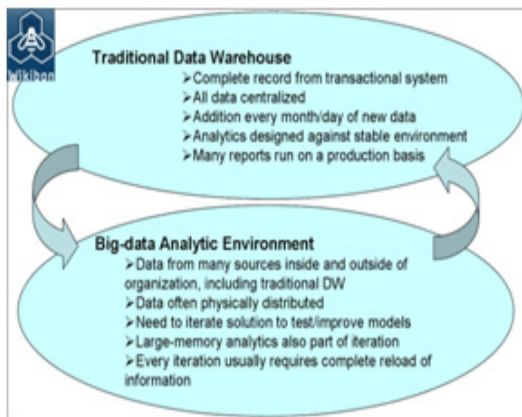


Fig. 3:

to be able to handle data warehouse. Traditional, reliable and efficient analytical tools they will need, video, etc. This type of data is growing at a very high speed.

The database requirements are very different in terms of large Data. Analysis of large amount of data is that can be used anywhere and in large quantities. Analysis of large data provides useful information. Hidden patterns are discovered. It focuses on unstructured data. Some technologies like Hadoop, NoSQL and Analytics are needed to map large data analysis data. Hadoop system is to capture data from different sources and then act in largest reduction. Thus, cleaning, storage, distribution, indexing, conversion, search, access, analyze and visualize as unstructured data into structured data is converted.

Figure 3 & 4 is shown a difference between traditional and big data analytics.

The operating principle behind Hadoop and big data for data query, the query processor to process the data does not have to move. Big data analysis used in different languages, Java, JavaScript, etc. problem based on different approaches to the analysis of traditional or advanced, Oracle Big Data needs to be done. It depends on the particular type of problem. Some traditional data warehouse analysis requires some more advanced techniques, including concept. But,

#	Dimension	Traditional	Big Data
1	Data Source	Mainly internal	Both inside and outside organization including traditional data warehouse
2	Data Structure	Pre-defined structure	Unstructured in nature
3	Data Relationship	By default, stable and interrelationship	Unknown relationship
4	Data Location	Centralized	Physically highly distributed
5	Data Analysis	After the complete build	Intermediate analysis, as you go
6	Data Reporting	Mostly canned with limited & pre-defined interaction paths	Reporting in all possible direction across the data in real time mode
7	Cost Factor	Specialized high end hardware and software	Inexpensive commodity boxes in cluster mode
8	CAP Theorem	Consistency - Top priority	Availability - Top priority

Fig. 4:

IT technology and tools requires for run large data processing, again very important and exciting. Big Data technologies are storage traditional data to work faster.

#### Technologies used for big data analytics

A NoSQL database can handle data stored in structured and unexpected data. The NoSQL database is generally high quality. NoSQL database storage and relational databases is used in a tabular modeling different relationships, in the instrument provides a mechanism for restoring data. Relational and NoSQL data model are very different. Relational data model takes the rows and columns separating it into several interrelated tables.

But NoSQL database oriented document takes the data in JSON format using documents. JSON is an important difference JavaScript notation. Another product that is relational model NoSQL, while is schema less Technologies rigid scheme. Many NoSQL databases is an excellent storage capacity integrated cache. Thus, the data used frequently. The system is placed in the database memory. NoSQL are:

- 1) Document Database: each pair of keys with complex data structure known as documents. Documents can include nested documents. This type of database to store (XML) semi-structured documents that are

generally hierarchical in nature unstructured (text) or.

- 2) Graphic reservations: Graphic database is based on graph theory. To store this information on the network is used.
- 3) The key stores of value: each item is stored as well as the name explains its value.
- 4) Wide column stores, that are customized to more questions instead of a column with large data sets and data storage Lines. Hadoop data processing is fast growing large Platform open source software. Hadoop can handle all types of Structured, unstructured, such as image or audio data. Its runs on Linux, OS / X, Windows and Solaris. Scalable, flexible and guilt are tolerant. It Hadoop HDFS.

File system scalable Hadoop HDFS distributed written in Java. This is explains HDFS architecture. HDFS master / slave architecture. A cluster consists HDFS Individual Name Node. Namespace manages the file system.

#### Challenges and opportunities with big data

Big data analytics face different challenges. It's It is described as follows:

- 1) The diversity and incompleteness: The machine algorithms expected homogeneous assays Data, and you cannot understand the



Fig. 5: Is shown a values and difficulties of big data storage and analysis for business model in future

nuances. Even after data Cleaning and bug fixes, some incomplete and the data is likely to be some mistakes [1].

- 2) The timeliness there is many situations in which the result of the analysis is required immediately. Given a Large data sets, it is often necessary to find the elements large data sets that meet specified criteria to process, analyze the time it will take. That is Data is growing difficult to design a structure Very high speed.
- 3) Human cooperation: a large data analysis system many of entry should support human experts, and shared the results of exploration.
- 4) Privacy and security: it is another great challenge Protection of personal privacy. E.g. the health industry, the history of the person is Personal. But it may be available from many sources. Therefore, it is difficult to maintain privacy and security.
- 5) Quality of data: a large amount of data is processed. The data are analyzed and it is important to capture the big challenge.
- 6) Analysis: Large data from different data sources. Analytics is a challenge.
- 7) Skills: Big data requires people with new skills. Management of large data effectively requires accurate people.

#### **Opportunities are summarized with large data**

- 1) Access to large data will become an important base Competition and growth of individual firms. All The companies will use big data. In most industries, Established competitors and new entrants will be equal Data-driven to innovate, to take advantage of competitive strategies; Deep and updated in real time and the value of the catch Information<sup>10</sup>.
- 2) Opportunities and Importance of big data in education. More detailed information can be generated. This for school teachers and parents is beneficial.
- 3) About the computer and electronic products, insurance, and government in all areas will

increase as their big data using productivity of large data.

- 4) Concept has practical applications in agricultural environment research. In agricultural crop care, seeds records, land records, environment records data, radiology images of crops and water supplying and environment effect on crops, and more information are analyzed coming from seeds and crop genetics and crop damages in relation to agriculture care. So studies can be completed faster. Big data will help to improve future diagnosis and treatment of the crops. Smart phones and tablet are use for watch and analyze of environment effect on crops.
- 5) The use of high volumes of mobile data traffic occurs. Big data is important for mobile networks. The network quality, traffic planning, hardware maintenance, etc. is useful to improve forecasting.
- 6) Generates a large amount of different branches of science experimental data. To meet the demands of science handle data requires a new kind of<sup>11</sup>.

Big data is a relatively new phenomenon. Large areas of data are different way of being for search of valuable business information. Big Data Analytics to analyze unstructured data, providing new ways for businesses and government. Research and development in this field is essential. There are several technical challenges that must be addressed. Research is needed to find a new way of handling data. Many IT decision makers, big data analytics tools and technologies is now a top priority.

Big data is playing a very important role in the future, which can handle big analytical software does not need big data. To requirement more storage as well as processing to extract new value, will be focused on the effective Analytics Growth. Technologies and tools for big data and continues to evolve.

## REFERENCES

1. Payal Malik, Lipika Bose," Study and Comparison of Big Data with Relational Approach", *International Journal of Advanced Research in Computer Science and Software Engineering*, **3**(8): pp 564-570 (2013).
2. Wei Fan, Albert Bifet," Mining Big Data: Current Status, and Forecast to the Future", *SIGKDD Explorations*, **14**(2).
3. Srinivasan,"SOA and WOA Article, Traditional vs. Big Data Analytics, "Why big data analytics is important to enterprises"
4. Elena Geanina ULARU, Florina Camelia PUICAN, Anca APOSTU, Manole Velicanu," Perspectives on Big Data and Big Data Analytics", *Database Systems Journal*, **III**, no. 4/2012
5. Big Data Survey Research Brief", Tech. Rep. SAS, (2013)
6. Dhruba Borthakur, The Hadoop Distributed File System: Architecture and Design
7. NESSI-Big Data White Paper," Big Data –a new world of opportunities" December 2012
8. Singh S. 2008. Agricultural Mechanization Policy. Proceedings of Tractor & Farm Machinery Manufacturers' Meet, Nov. 16-17, 2007, CIAE, Bhopal.
9. Kulkarni SD. 2005. Food Safety and Security Issues in India: Challenges and Approach. Lead Paper for presentation in National Seminar on Post Production Systems and Strategies to the Issues and Challenges of Food Safety and Security during Sept. 22-23, 2005 at TNAU, Coimbatore – 641003
10. Ali, N. 2004. Rural development in India through post-harvest technology and value addition activities in the agricultural production catchment. Paper presented at the International Conference on Emerging Technologies in Agricultural and Food Engineering to be Held at IIT, Kharagpur during 14-17 Dec., 2004.
11. Bhattacharrya, P. and G. Chakraborty. Current Status of Organic Farming in India and other Countries. *Indian Journal of Fertilisers*, **1**(9); (2005).
12. Kapil Shukla, deven patel, bankim radadiya, "Role of Information Technology in Improvement of Current Scenario in Agriculture", Paper presented at the Oriental Scientific Publishing Co., India, Vol. 7, No. (3): Pgs. 390-395.