

Conceptual paper on Application of Big data in Risk Management

Chethan Jayantha¹,

Research Scholar, Research and Development Centre,
Bharathiar University, Coimbatore 641 046

Dr. A. Dhanalakshmi²

Research Supervisor, Research and Development Centre,
Bharathiar University, Coimbatore 641 046

Abstract

Big data has become one of the most important technology trends that has the potential for dramatically changing the way organizations use information. Big data enables organizations to gather, store, manage, and manipulate vast amounts of data at the right speed and time, to gain the right insights. Organizations across industries and geographies have started harnessing the revolutionary potential of the Big data technology. Risk management is an area in which the potential of Big data is yet to be ventured. This paper gives an insight into the Big data Technology and process of integration of Big data in the different stages of Risk management.

Key words

Enterprise Risk Management (ERM), Big Data

Introduction

With the advent of Information technology revolution the quantity of data getting accumulated is enormous. As we are reading the article, the world's data is exploding in unprecedented velocity, variety, and volume. Data is getting piled up from every aspect of our modern day life. Starting from purchases made in stores to the posts in social media sites. 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 (*IBM, Bringing big data to the enterprise*). This is big data. The Big data is a gold mine and it is up to the corporates to capitalize this. Big data's potential has already been harnessed in different areas of management. Risk management is a potential area in where the big data can be embraced.

The structure of the paper is organized into 3 Segments. The first segment concentrates on understanding Big data, its need and process of harnessing the Big Data. Second segment is understanding the Risk and Risk management process. The final segment talks about the application of Big data in Risk management.

Segment 1 :BIG DATA

Need for Information:

The success of an organization depends on the decisions which its management takes. The well informed decision which is taken on a timely basis, will most likely lead the organization to the success. Hence the main requirement of the management is to get correct information on a timely basis. This forced the management to collate both financial and non-financial information from various sources to get correct and timely information. Historically it was a simple task. Company had only a handful of customers and very few products and things were pretty

straightforward and simple. The decision used to be taken based on the historical information accumulated by the organization namely financial statements, credit history of the customer etc., Separate database was created for each data sources namely financial records, customer master data, customer credit records, employee records, inventory records to name a few. Since the transaction were very minimal there was no need to integrate this data bases.

But over time, companies, its portfolio and the markets have grown and the complexities have grown multifold. Now the focus of information has shifted from completely relying on the internal data to the integrating the internal data with the external data and take a pragmatic approach in decision making. The greatest challenge today is to integrate all these different types of data. Hence, organizations are finding ways to make use of this big data. This is the reason for evolution of Big Data Technology and Big Data Management.

Definition of Big data:

In 2012, Gartner defined Big Data as follows:

"Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization." Additionally, a new V "Veracity" is added by some organizations to describe it. This definition is widely accepted and referred to as Four V's definition of Big data.

According to this definition following are the characteristics of the Big Data are:

- Volume
- Velocity
- Variety
- Veracity

1. Volume

As the name of Big Data indicates, the size of the data is Big. In other words it refers to the vast amount of data which is getting generated every second. It is not just Gigabytes or Terabyte it is Zettabyte and Brontobytes. It is estimated that 2.5 Quintillion Bytes (i.e. 2.3 Trillion Gigabytes) of data are created every day. Further, on an average most companies are storing 100 terabytes of data. It is also estimated that 40 Zettabytes (i.e. 43 Trillion Gigabytes) of data will be created by 2020.

2. Velocity

High Velocity refers to the speed in which the data is getting generated and getting transmitted. Posting in the social media which goes viral in second is a good example of the High Velocity of the data. It is estimated that New York Stock exchange captures 1 TB of trade information every day. Further, it is predicted that by 2016, there will be 18.9 Billion network connection it almost 2.5 connections per person on earth.

3. Variety

High Variety refers to the different type of data getting generated. Historically the type of data getting generated was restricted to the Alfa-numerical i.e. text and numbers. These data could be easily mapped using tables and rational databases. In current scenario the data which is getting generated are unstructured in nature. Few example are images, video, voice, signals, GPS tracking etc. It is estimated that unstructured data constitutes over 80% of the total data being generated. On an average 30 Billion pieces of contents are shared in Facebook every month. 400 million tweets are sent per day by about 200 million active users.

4. Veracity

Veracity refers to the quality of the data which is getting captured. In other words it is the uncertainty of the quality of the data getting captured. It is estimated that poor data quality cost around \$ 3.1 trillion a year in US alone.

Harnessing the power of Big Data

GIGO (Garbage In Garbage Out) is the standard rule in the Data management and Data processing. It means that if garbage is feed for data processing then the Garbage is generated as end product. It is an exaggeration of the fact that if we need a meaningful and useful information out of data processing then the processing should be done on the correct data. Big Data is a huge junk of data, it is up to the user to get a productive output from it. Following steps need to be followed in this regard.

a. Determine the objective and goal

The first step in Big data management is to identify the organizations objective and goal. The organization has to list down the things what they need from the big data. In other words, in big data management it should always be begin with the End objective. If the organization is not determined its objective of using the big data then it will be lost in the processing the Big data without any productive output.

b. Establish your road map

The next step in the Big data management is to put the road map to achieve the objective and goal set by the organization. In this stage the process and activity which need to be put in place is discussed with all the stakeholders and agreed upon. It will act as a benchmark for the Big data management.

c. Discover your data

In this stage the data which is required for achieving the desired objective is identified. When the required data is identified it would be easy for the organization to mine the Big data to the get the desired data. For example if the organization is analyzing the online spend pattern of customer then the data requirement would be credit card transaction, debit card transaction, merchant bank transactions etc.,. When this is identified then it will be easy of the organization to segregate these type of transaction with other transaction in junk of Big Data.

d. Figure out what data you don't have

One of the basic characteristics of Big data is Veracity. Therefore, it is most likely that the data extracted will be of poor quality. Hence the next step is to identify the missing data. There are various scientific tools available to overcome this drawback of big tools. These tools will extrapolate the missing data based on the available data, there by complete the loop of missing data.

e. Understand the technological option

The next important step is to identify the technological solution available for processing the Big data. It is a very vital step in the Big data management as it involves a huge cash outflow in terms of purchase of necessary hard ware and software to support the Big data processing. There are mainly option available in terms of software solutions from various service providers. The organization has to wisely select the option considering the objective set and the cost benefit analysis.

f. Continuously monitor and test assumptions

After the desired result is achieved from Big data, the next step is to continuously monitor the activity of Big data processing and continuously revisit the assumption of the required data. It is vital as new variety of data is getting pumped in on a regular basis and the data which was important earlier may be redundant in the later point of time.

Segment 2: Risk Management

There are two basic parameters which governs the business decision namely Risk and return. Every Business is exposed to some or the other risk, which affects its profitability and cash flows. Risk is a probability or threat of damage, injury, liability, loss, or any other negative occurrence that is caused by external or internal vulnerabilities, and that may be avoided through preemptive action (*Business Dictionary*). In broad terms the uncertainty associated with the outcome of an event that can lead to loss. Risk management is a technique by which enterprises manages to establish a tradeoff between the two parameters for minimizing the impact of risk and maximizing the stakeholders wealth.

ISO 31000 defines Risk management as follows:

Risk management is a process of identification, assessment, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events or to maximize the realization of opportunities. Risk management's objective is to assure uncertainty does not deflect the endeavor from the business goals

Process of risk management

1. Identification of Risk

The Risk management process starts with the objective setting and identification of Risks which will impair the organization from achieving the objectives. Objective of the organization are the mission or vision which the management sets and cascades the same throughout the organization. Organization aligns all its activities to these objectives. Any activity which affects the organization from achieving its objectives are termed as Risk.

Following are various methods of risk identification:

- Objectives-based risk identification – This is the common and widely used method of risk identification. In this method all the activities which has the potential of endangering the achieving an objective partly or completely is identified.
- Scenario-based risk identification - In scenario analysis different scenarios are created. The scenarios may be the alternative ways to achieve an objective. Any event that triggers an undesired scenario alternative is identified as risk.
- Taxonomy-based risk identification - The taxonomy in taxonomy-based risk identification is a breakdown of possible risk sources. Based on the taxonomy and knowledge of best practices, a questionnaire is compiled. The answers to the questions reveal risks.

- Risk charting–In this method all known risks are listed and a matrix is prepared as to which resources are affected by the said risk (*Crockford, Neil, 1986*).

2. Assessment of Risk

After the risk is identified the next step is to identify the Probability and severity of risk. Probability of risk is the likelihood and frequency of occurrence of risk. For example the probability of earth quake with 7.5 magnitude in Richter scale is less whereas the probability of road accidents are high.

Severity of the risk is the impact of the risk. In other words it is the extent of loss the organization will face when the risk event has occurred. For example the impact of earth quake with 7.5 magnitude in Richter scale is very high whereas the impact of theft in the storeroom is low.

Based on the likelihood and impact of Risk, the Risk can be classified into 4 quadrants. Namely

- a. High probability and High Impact
- b. High probability and low Impact
- c. Low probability and High Impact
- d. Low probability and low Impact

3. Risk Response

After the Risk assessment is done the next step is to determine the appropriate risk response. Risk response of the organization depends based on the Risk Appetite of the organization. Risk Appetite is the level of risk that an organization is willing to accept. However as a thumb rule following Risk response is adopted.

- a. Risk avoiding: In this case the entire activity is avoided to avoid the risk from the activity. This response is generally adopted when the probability of risk as well as the impact of risk is high. For example: A manufacturing activity of an organization involves handling of radioactive material which may have harmful impact on the employees. Here the likelihood of risk is high and impact is also high. In this case the organization may decide to stop this manufacturing activity.
- b. Risk reduction: When the probability of risk is high and the impact of the risk is low, then the organization would prefer to adopt the risk reduction process. In risk reduction process the activity which will increase the probability of risk will be identified and reduced.
- c. Risk Transfer: This method of risk response is adopted when the likelihood of risk is low and the impact of the risk is high. In case of risk transfer the risk of the organization is transferred to other parties. The best examples of risk transfer is Insurance and outsourcing. In case of insurance the risk of the organization is transferred to the Insurer/Insurance Company for a consideration called premium. In case any fatality occurs the organization suffers loss then the loss will be made good by the Insurer/Insurance Company.
- d. Risk Retention: When the probability of risk and impact of risk is low then the best option is to retain the risk. In other words the said activity which has risk is continued.

4. Implement of Risk response and monitoring

After the risk response is finalized the next step is to implement the same. The process of Risk management will not end with the identification and implementation of Risk response. It is a continuous process. After the implementation of Risk response next step is the continuous monitoring of the implantation status of the Risk response and to identify the new risk and the process of Risk management flows on.

Segment 3: Integration of Big data technology in risk management

Different risk domains stands to benefit from Big Data Technologies in diverse ways. Big Data can be targeted to organization's particular needs whether they are greater volume, variety or veracity and strategically applied to enhance different risk domains. This segment of the paper is to understand the way of integrating the Big Data in Risk management.

1. Identification of Risk:

The biggest limitation of the existing architecture of the Risk management is that it cannot be scale efficiently to meet the increasing types of risk. Risk which the organization faces in the current scenario is entirely different from the traditional risk. To name a few, Reputation risk, Information technology related risk namely hacking, virus etc., Now the focus is more on the external risk which is uncontrollable than on the internal and controllable risk. The need now is to identify these risk on a real time basis. BigData technologies present fresh opportunities to address these challenges. Vast amount of data, which is comprehensive enough and near real-time data has the potential to improve identification of risk.

2. Assessment of Risk

Assessment of Risk stands on two pillar namely probability of risk and severity of risk. Big Data technology can be used for assessment of both these parameters of risk. Big data technology has the vast data with vast timeframe. This details can be used to identify the number of instances when the risks have matured in the past. These details can be extrapolated using statistical tools to identify the probabilities of occurrence of the said risk in the future. Similarly the severity of the risk which had matured in the past can be measured and indexed. This information can be used as benchmark for measuring the severity of the risk which may occur in future.

3. Risk Response

Big Data technologies will allow the development of models that will support the organization in finalizing the risk response. Big data technology has the ability to process enormous amounts of data in fast timeframes and accommodate new requirements for scenario stress tests at the trade, counterparty and portfolio levels. With this traits of the Big data technology, Scenario simulations models can be built which serves in identification of cause and effect relationship. This model in collaboration with the heat map can be used to test the different risk response scenarios. Thereby giving an opportunity to the organization to test the risk response in the test environment before implement the same.

4. Continuous monitoring

Big data technology has the potential of near real time data. It can be configured to generate sophisticated reports which has the power of predicting the occurrence of risk on a real

time basis. This facilitates the organization in monitoring of risk not restricting to the internal risk within the organization but also to the external risk which is universal in nature.

Conclusion:

There is tremendous scope and opportunity to leverage Big Data technologies to improve riskmanagement systems. This paper examined various aspects of riskmanagement and how they can be integrated with the Big Data technologies. Although Big data opens new avenues in risk management, the organization has to perform a cost benefit analyses and decide the implementing of integrating the Big Data technology in Risk management to have an efficient and efficient organization.

Bibliography:

1. Business Dictionary <http://www.businessdictionary.com/definition/risk.html#ixzz3pl6BfMHB>
2. Crockford, Neil (1986). *An Introduction to Risk Management* (2 ed.). Cambridge, UK: Woodhead-Faulkner. p. 18. ISBN 0-85941-332-2.
3. IBM, Bringing big data to the enterprise <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>
4. ISO/IEC Guide 73:2009 (2009). *Risk management — Vocabulary*. International Organization for Standardization.
5. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, and Marcia Kaufman(2013) , *Big data for Dummies*, ISBN 1118644174