JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BIG DATA AND ITS VARIOUS PRIVACY ISSUES

[1]Mrs.M.Angel shalini,[2]Ms.H.Linda Evangelin,
[1] Assistant professor, [2]Resreach Scholar,
[1]PG And Research Department of Computer Applications,
[1]Hindusthan College of Arts And Science, Coimbatore, Tamilnadu, India.

Abstract- Big data can be defined simply as huge data sets that outgrow databases and data handling. In data mining the process of going through a large set of data to find relevant information. The information was collected from various sources and a wide range of useable fields. The dissemination of data by scientifically developed models. Big data conventional methods are having structured and unstructured information. In the future, big data will be very useful for businesses and society like having a network facility like telecom network facilities. Previously there were building estimates and virtual models of reality now that can be done using already collected information from various sources. Big data is now almost as become a part of our society like communication, marketing, research, banking, transportation, defense fields. Big data is opened possibilities of use in the science health care system, economic decision, educational field, etc, where public interaction and amusement. But these opportunities also compromise our privacy and security concerns. Big data always uses information from various sources of data like the cloud and sometimes it may require additional distributional servers to complete its job. It was also observed that the process of big data also compromises our security concerns. This edition of big data will focus on the security concerns related to it.

1. Introduction

The word "big data" is a term used to describe a great volume of information processed and unprocessed. As the information collected is very huge, it is extremely difficult to use conventional processing techniques. In many institutions, the information collected is too large or it may be moving at extremely high speed [1]. Where there will be unable to process with conventional data processing machines. With the advent of big data, businesses may improve their operation capabilities by using big data as a tool to improve their business performance quickly. The usual concept of big data appears to be huge but that is not a situation in reality [2]. Businesses, companies usually refer to big data as a technological tool and in processing big data storage devices.

Big data is an efficient tool in the management of next-generation information technology industries. Some of the industries based on the internet of things, cloud computing and big data can solve many deliberate problems. Even though data warehouses are used to manage datasets it is a long and time-consuming process in extracting valuable information from big data. Data mining may not be enough to handle large datasets. The important problem in the analysis of big data is the lack of coordination between different databases

tools such as data mining and statistical analysis. It is a difficult proposition in a data mining situation [2]. The insistence of data for usable applications is a major ultimatum. There is a clear implication in describing data revolutions. In general, it is assumed that big data was initiated by big companies who are handling network data storage devices and look for possibilities of distribution of collected data efficiently. Usually, the information collected from various sources of data collection of devices would be extremely huge of data consisting of the huge number of records from millions of people which is related to health care, business governance, transportation, etc, usually, a type of information collected would be unprocessed and ineligible to be approached. It is not necessarily referred to data as a huge information tool and it leads to the possibility of the process of finding new ways of managing the big data.

2. Characteristics of big data

Big data usually comprises a huge amount of information which existing software techniques are insufficient and inefficient in managing data structures. There are different methods of explaining big data [3]. It is a term generally used to explain big data like velocity, variety, volume, veracity, and value. Volume is a piece of information collected in bytes; velocity depends upon the speed of the processor; variety is various sources of data; veracity indicates the quality of the data and reliability of the data; value usually denotes the large information made valuable.

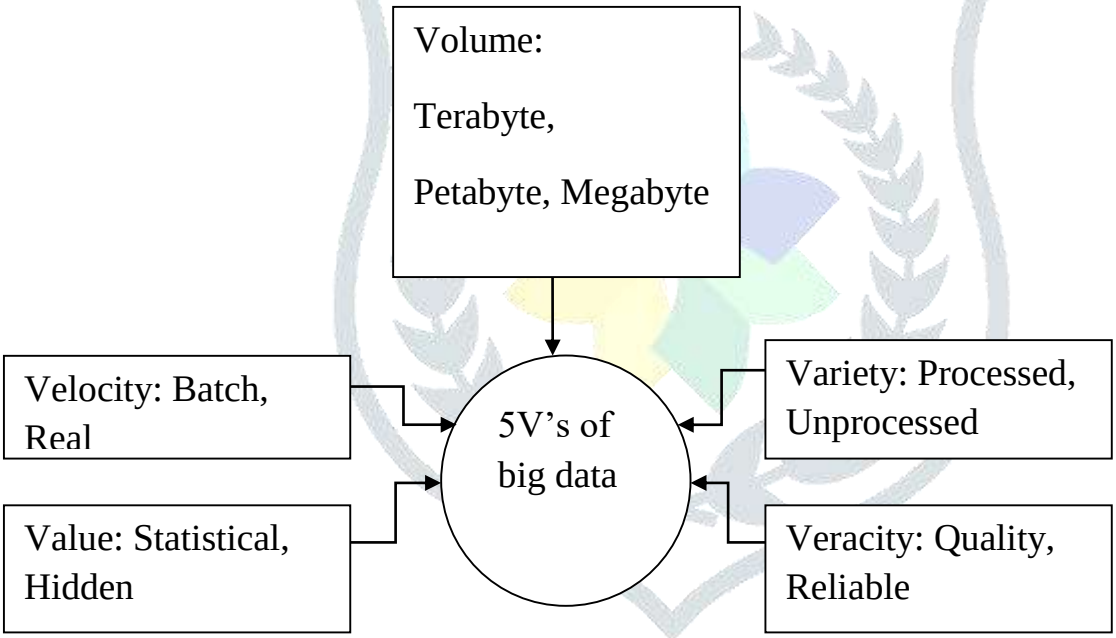

Figure 1: Characteristics of big data

2.1 Volume

Volume can be defined as a scale to measure the amount of data generated every second like cell phones, cars, sensors. The huge amount of data generated cannot be analyzed by using existing methods [3]. Now we use distributed system, which is stored at different locations and brought together by software with Facebook alone there are 10 billion messages, 4.5 billion times like buttons are pressed and millions of pictures are uploaded every day By collecting and analyzing these data is a great challenge of large proportions.

2.2 Velocity

Velocity is the speed limit in which the data is being generated, collected, and analyzed. Every day there is an increase in the number of emails, Twitter messages, images, videos at a lightning speed in the world. Every second a data is increasing [3]. Not only it must be analyzed, but the transmission speed and access to the data also remain immediately to allow for real-time access to the website, credit card verification. This big data technology allows analyzing the data which is being generated, without even putting it into databases.

2.3 Value

When we analyze data for extracting there is a definite value. It may have useless data which is not applicable in our everyday life. Unless there is useful data imagine there is a clear link between data and insights, so this does not always mean there is value in big data. in big data, there should be an initiative to understand the costs and advantages of collecting and analyzing the data to make sure that the data is reaped can be monetized.

2.4 Variety

Variety is referred to as the different types of data. Data from the past looks very different from the present. Previously we had structured data that fits well and neatly into a data table. But present data are unstructured. Generally, 80% of the world's data accommodate into this category which includes image, video, social media updates, etc [3]. Big data technology is now allowing structured and unstructured data to be reaped, stored, and used simultaneously.

2.5 Veracity

Veracity is the quality or trustworthiness of the data. The accuracy of data cannot be determined. For example, in a Twitter account, there will be a lot of hashtags, abbreviations, typos, and the reliability and accuracy of that content. Cleaning millions of data is no use if the quality of the data is not accurate. This relates to the use of global positioning system data. Satellite signals are lost as they jump tall buildings or other structures. If this happens, the location data has to be fused with another data source like road data

3. Classification of big data

Locating the hidden patterns may help extract information from big data. Big data is divided into ten divisions as data format, data type, data source, data consumer, data store, data usage, data analysis, data processing propose, data frequency, and data processing method.

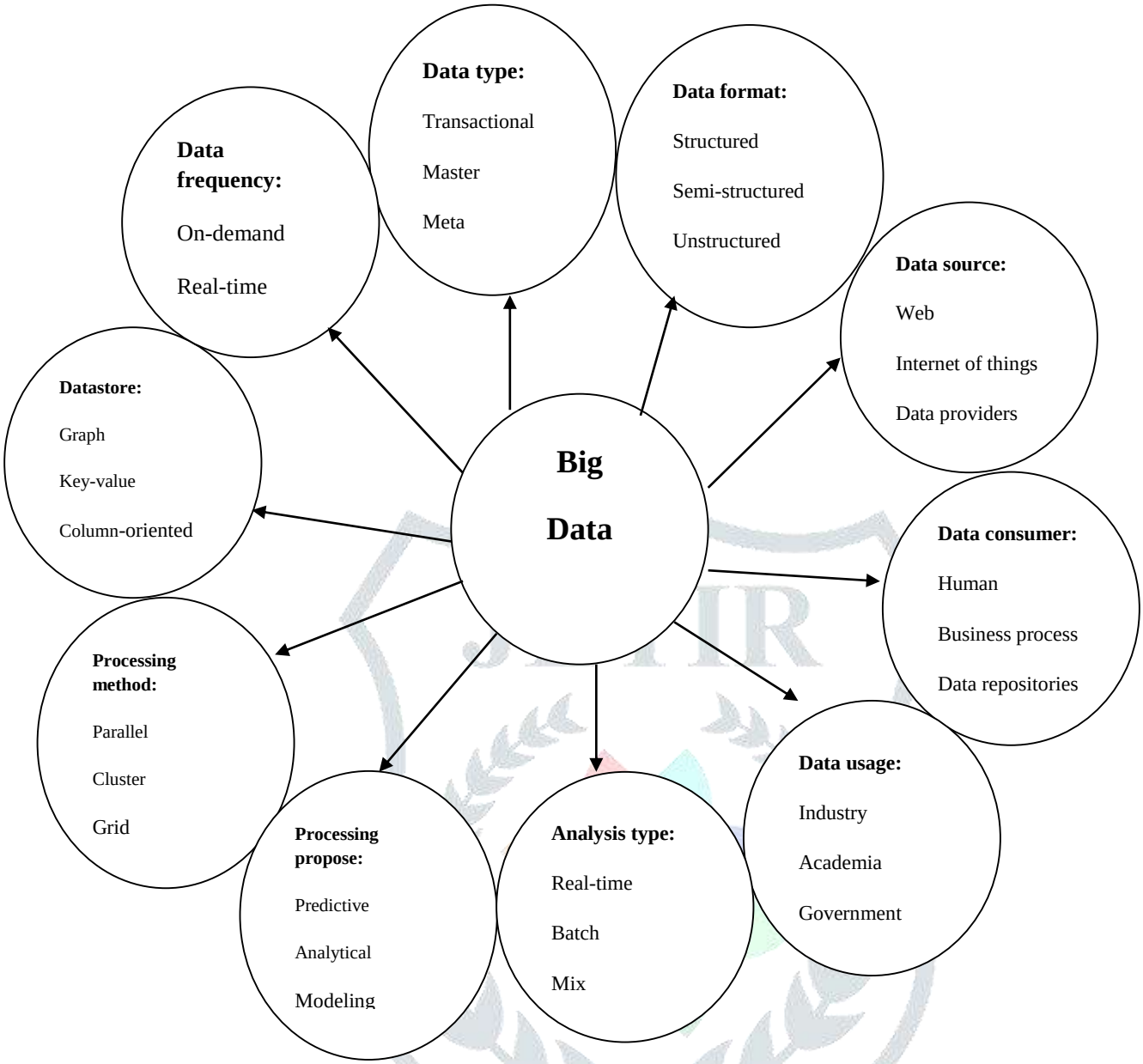

Figure 2: Classification of big data.

4. Stages involved in big data

4.1 Data Acquisition

In Big Data, the data acquiring itself is the first step. The technology world is growing and data generation is rising exponentially. Introducing smart devices which are used with a wide array of sensors generates data continuously. The Large Hadron Collider in Switzerland produces petabytes of data [4]. Some of the data is not useful and can be abandoned, however, due to its unstructured form; selectively discarding the data presents a dispute. This data becomes more powerful when it's combined with other valuable data and insisted. Due to the interconnection of devices over the World Wide Web, data is increasingly being merged and stored in the cloud.

4.2 Data Extraction

The data which is generated and acquired is not of use. It contains a massive amount of invalid or unimportant data. For instance, a simple CCTV camera uniformly polls the sensor to collect information about the user's action. However, when the user is in a state of redundant, the data generated by the activity sensor is also redundant and of no use. The problems presented in data extraction are twofold: firstly, due to the nature of data generation, deciding which data to keep and which to discard data increasingly depends on the context in which the data was initially generated [4]. For instance, footage of a security camera with the same frames may be abandoned however it is important not to discard the same data in a case where it is being generated by a heart-rate sensor. Secondly, there is a lack of a common platform that has its own set of challenges. Due to the large variety of data that exists, bringing them under one platform to standardize the data extraction is a major challenge.

4.3 Data Collation

Data from a single source often is not enough for investigation or prediction. More than one data source is often combined to give a large sizeable picture to analyze. For example, a health monitor application often gathers data from the heart-rate sensor, pedometer, etc. to abridge the health information of the user[5]. Likewise, weather forecast software takes in data from many sources which reveal the daily humidness, temperature, precipitation, etc. In the scheme of Big Data convergence of data to form a bigger picture is often considered a very important part of organizing.

4.4 Data Structuring

Once all the data is collected, it is very important to present and store data for further use in an ordered format. Organizing is important so authoritative information can be made on the data. Data structuring employs methods of arranging the data in a particular model. Various new platforms, such as Non-SQL, can query even on unordered data and are being increasingly used for Big Data Analysis. A major issue with big data is providing actual-time results and therefore structuring of collected data needs to be done at a rapid pace.

4.5 Data Visualization

Once the data is ordered, the queries are made on the data and the data is presented in an optical format. Data Analysis involves picking out areas of interest and providing results based on the data that has been structured. For instance, data containing average temperatures are shown close to the side water consumption rates to calculate a relation in between them[5]. This examination and presentation of data make it ready for consumption for users. Raw data cannot be used to gain understanding or for judging patterns, therefore “civilizing” the data becomes all the more important.

5. Applications of big data

In a current situation, the companies would not miss the big data opportunities because the innovation, competition, and productivity of the company will go down. Big data is a tool and technology that helps companies to interpret a large amount of data very faster. It helps boost production efficiency and also to develop new data-driven products and services[6]. The applications are:

- Retail/consumer
- Finance & Frauds services
- Web and digital media
- Healthcare & life science
- Telecommunications

- Ecommerce & customers services

6. Issues in big data

Big data is useful as well as important that how made efforts to tap into as well, so it can be used for its evil purpose. There are some issues with big data. They are,

- Privacy breach
- Impossible to anonymous
- Analytical isn't accurate
- E-discovery problems
- Personal information at risk
- Discrimination
- Legal protections are only a few for involved individuals
- Patent and copyright may be irrelevant in the future
- Interpretations can become ethical issues
- Big data is here to stay
- Vulnerability to fake data generation
- The potential presence of untrusted mappers
- Troubles of cryptographic protection
- Possibility of sensitive information mining
- Struggles of granular access control
- Data provenance difficulties
- High speed of Non-SQL databases evolution and lack of security focus
- Absent security audits
- Analytical problem with data
- Data theft is evident

6.1 Privacy breach

Data is today's commercial currency. So, it stands to reason that criminals today will use every means necessary to breach your security to access your data. For your organization to be safeguarded from a data breach, you will need aspects of understanding of the types of data breaches or attack vectors available to the cyber lawbreaker.

6.1.1 Hacking

Cyber offenders are getting smarter day-to-day and are continuously using a diversity of techniques both new (zero-day) as well as variations on old exploits.

6.1.2 Insider danger

Your employees know the most about where you're most careful data exists and, in some cases, how it is protected, so they can impose significant damage if not properly monitored or security protocols put in place.

6.1.3 Data on the move

We live in a progressively mobile world, so another concern has to be when laptops or flash drives are stolen, or backup tapes are off-track in the mail.

6.1.4 Physical theft

Although having Ethan Hunt fly down an air vent to physically access a fixed network is a thing of Hollywood lore, physical robbery is a reality. Perhaps, not as dramatic as in film, physical data theft can be as simple as plugging a USB drive into a sensitive.

6.1.5 Improper disposal

People make an error all of the time, so it is expected that at some point someone will do something dumb when it comes to data handling.

6.1.6 Unauthorized access

This form of a data breach is directly assigned to an absence of access controls. Specifically, if admin benefits are poorly monitored or there are no controls of the level of privilege within specific applications or even across network resources.

6.2 Personal information at risk

As more personal details are collected up by ever-more-powerful computers, giant sets of data – big data – have become available for not only lawful uses but also abuses.

Big data has a vast potential to transform our lives with its predictive power [7]. Visualize a future in which you know what your weather will be like with 95 percent accuracy 48 hours ahead of time. But due to the likelihood of harmful use, there are both security and privacy threats of big data you should be worried about, especially as you spend more time on the Internet.

6.2.1 The size of the potential problem

First of all, due to the complete scale of people involved in big data security events, the poles are higher than ever. When the professional development system at Arkansas University was ruptured in 2014, just 50,000 people were pretentious. That's a large number, but compare it with 145 million people whose birth dates, home and email addresses, and other data were stolen in a data breach at eBay that same year. From the outlook of a security professional, protecting big data sets is also more daunting. This is partly due to the nature of the underlying technologies used to store and process the details. Big data companies like Amazon heavily depend on distributed computing, which typically involves data centers geographically scattered across the whole world. Amazon divides its global operations into 12 regions each containing multiple data centers and being potentially subject to both physical attacks and tenacious cyberattacks against the tens of thousands of individual servers housed inside.

6.2.2 Difficulties with access control

One of the best master plans for direct access to information or physical space is having a single access point, which is much easier to secure than hundreds of them. The fact that big data is stored in such widely increased places runs against this truth. Its vulnerability is far higher because of its size, distribution, and wide range of access. In addition, many enlightened software components do not take security seriously enough, including parts of companies' big data infrastructure. This opens a further direction of potential attack. For instance, Hadoop is a collection of software parts that allows programmers to operate a large amount of data in a dispensed computing infrastructure. When first introduced, Hadoop had characteristics suitable for a system used by only a few users. Many big companies have adopted Hadoop as their corporate data platform, even though its access control mechanism wasn't designed for large-scale assumptions.

6.2.3 Consumer demand drives security and privacy

For consumers, then, it is disapproving to demand an intensified level of security through vehicles such as terms and conditions, service level concurrence, and security trust seals from organizations collecting and using big data. What can companies do to protect personal information? Countermeasures such as

encryption, access control, intrusion detection, backups, auditing, and corporate methods can prevent data from being breached and dropping into the wrong hands. As such, security can encourage your privacy.

At the same time, enhancing security can also hurt your privacy: it can provide lawful excuses to collect more private data such as employees' web surfing history on work computers. When law imposition agencies collect information in the name of better security, everyone is treated as a possible criminal or terrorist, whose information may finally be used against them. The decision-makers already know a lot about us but could ask companies such as Apple, Google, and Amazon to provide more intelligence such as a decrypted version of our data, what search terms we are using and what we are buying online.

The basic security principle used to justify this type of blanket observation (which is now more affordable and feasible due to the use of big data technologies) is "nobody can be faithful." Once the data is collected, those data can join the rest of the data in being receptive to abuse and breaches, as demonstrated in enquiring incidents involving National Security Agency employees.

And yet when used properly, big data can help increase your privacy by permitting more information to be used to maximum advantage and eventually improve the quality (especially, the accuracy) of intelligence on possible attacks and attackers in cyberspace. For example, in an ideal world, we don't have to worry about dishonest emails (also called phishing) because a big data analytics engine would be able to pick out dangerous emails with pinpoint accuracy.

6.3 Vulnerability to fake data generation

Before things that are said to all the operational security challenges of big data, we should mention the concerns of fake data generation. To intentionally undetermined the quality of your big data analysis, cybercriminals can forge data and 'pour' it into your data lake. For occasion, if your manufacturing company uses sensor data to identify malfunctioning production processes, cybercriminals can pierce your system and make your sensors show wrong results, wrong temperatures. This way, you can lose to notice alarming trends and miss the opportunity to solve problems before serious damage is caused. Such problems can be solved by applying the fraud detection method.

6.4 Potential presence of untrusted mappers

Once your big data is gathered, it experiences parallel processing. Map-reduce is one of the methods used here. When the data is split into numerous bulks, a mapper processes them and allocates them to particular storage options. If an outsider has access to your mappers' code, they can change the settings of the existing mappers or add 'alien' ones. This way, your data processing can be effectively ruined: cybercriminals can make mappers produce inadequate lists of key/value pairs. This is why the results brought up by the Reduce process will be faulty. Besides, outsiders can get access to sensitive information.

The problem here is that getting such access may not be too difficult since generally, big data technologies don't provide an additional security layer to protect data. They usually tend to rely on perimeter security systems. But if those are faulty, your big data becomes a low-hanging fruit.

6.5 Troubles of cryptographic protection

Although encryption is a well-known way of protecting sensitive information, it is further on our list of big data security issues. Despite the possibility to encrypt big data and the essentiality of doing so, this security measure is often ignored. Sensitive data is generally stored in the cloud without any encrypted protection. And the reason for acting so recklessly is simple. Which entails the loss of big data's initial advantage – speed?

6.6 Possibility of sensitive information mining

Perimeter-based security is typically used for big data protection. It means that all 'points of entry and exit' are secured. But what IT specialists do *inside* your system remains a mystery. Such a lack of control within your big data solution may let your corrupt IT specialists or evil business rivals mine unprotected data and sell it for their benefit. Your company, in its turn, can incur huge losses, if such information is connected with a new product/service launch, the company's financial operations, or users' personal information.
Here, data can be better protected by adding extra perimeters. Also, your system's security could benefit from anonymization. If somebody gets personal data of your users with absent names, addresses, and telephones, they can do practically no harm.

6.7 Struggles of granular access control

Sometimes, data items fall under restrictions and practically no users can see the secret info in them, like, personal information in medical records. But some parts of such items could theoretically be helpful for users with no access to the secret parts, say, for medical researchers. Nevertheless, all the useful contents are hidden from them. And this is where talk of granular access starts. Using that, people can access needed data sets but can view only the info they are allowed to see.

The trick is that in big data such access is difficult to grant and control simply because big data technologies aren't initially designed to do so. Generally, as a way out, the parts of needed data sets, that users have the right to see, are copied to a separate big data warehouse and provided to particular user groups as a new 'whole'. For medical research, for instance, only the medical information gets copied. Though, the volumes of your big data grow even faster this way. Other complex solutions to granular access issues can also adversely affect the system's performance and maintenance.

6.7 Data provenance difficulties

Data provenance – or historical records about your data – complicates matters even more. Since its job is to document the source of data and all manipulations performed with it, we can only imagine what a gigantic collection of metadata can be. Big data isn't small in the volume itself. And now picture that every data item it contains has detailed information about its origin and the ways it was influenced (which is difficult to get in the first place).

For now, data provenance is a broad big data concern. From a security perspective, it is crucial because:

- Unauthorized changes in metadata can lead you to the wrong data sets, which will make it difficult to find needed information.
- Untraceable data sources can be a huge impediment to finding the roots of security breaches and fake data generation cases

6.8 High speed of Non-SQL databases evolution and lack of security focus

This point may seem like a positive one, while it is a serious concern. Now NoSQL databases are a popular trend in big data science. And its popularity is exactly what causes problems.

Technically, Non-SQL databases are continuously being honed with new features. And just like we said at the beginning of this article, security is being mistreated and left in the background. It is universally hoped that the security of big data solutions will be provided externally. But rather often it is ignored even on that level.

6.9 Absent security audits

Big data security audits help companies gain awareness of their security gaps. And although it is advised to perform them regularly, this recommendation is rarely met in reality. Working with big data has enough challenges and concerns as it is, and an audit would only add to the list. Besides, the lack of time, resources, qualified personnel, or clarity in business-side security requirements makes such audits even more unrealistic

6.10 Analytical problem with data

Even though big data is changing businesses by providing actionable insights, there are certain problems related to it. A problem with big data is that it grows constantly and organizations often fail to capture the opportunities and extract actionable data. Companies often fail to recognize where they need to allocate their resources. This failure in allocating the resources results in not making the most of the information. Apart from that, organizations often end up with talent that does not understand how they should use big data analytics. Such a dearth of trained employees who can extract information results in companies not making the most of information held by them. Furthermore, while extracting insights from the big data held by them, companies fail to identify the right objective and end up with insights that are not so helpful for their growth.

6.11 Data theft is evident

When organizations store large amounts of data sets, these sets consist of almost every type of information that is even minutely meaningful for the company. As a result, when authorities fail to install proper security measures, they are susceptible to the threat of data theft. This information theft means that a company is losing out on vital information. Moreover, data theft can also disclose confidential information that the business has hidden over the years. This could mean a lethal blow to the business's reputation. Consumer information is the primary target for an attacker. By stealing such information from an organization, attackers can sell it to other companies for monetary benefits. Problems with big data are avoidable with proper solutions. CTOs and CIOs can start searching for ways through which they can avoid the problems of big data analytics. Securing big data is also an aspect that companies can take into consideration. Necessary changes in the infrastructure may be imperative to ensure that the data stays safe and usable...

7. Privacy issues on various applications

7.1 Health care system

Due to various types of diseases and advanced medical treatment in the history of patients is stored in a big data format. The information collected is used to identify the patients with high risks of health disorders at an early stage and provide improved quality care and reduce the cost of health care.

Even though there are many benefits, new research suggests that big data may be riskier than originally thought. As per the research conducted, it was found that the data collected from the patients is accessible to hacking. So it is very important to be secure about the implications of big data. Traditional security solutions are limited in scope to be applied for healthcare solutions because of diversity in formats and complexity in securing abnormally distributed software. That is the reason for incorporating big data analysis is necessary for exposing data to analysis.

7.1.1 Data governance

As the healthcare industry is moving towards a value-based industry healthcare analytics, data governance will be the primary object in regulating and maintaining healthcare data. The object is to have

common data that represent industry standards and local standards. Presently, this method of data generation by body sensor network is different and would require governance.

7.1.2 Privacy-preserving in healthcare

In the field of big data, intrusion into the privacy of patient data by other unauthorized persons is a big problem in big data. A report was published by a well-known magazine over patient privacy that created big concerns for the management. This incident made to consider privacy in big data. For example, data coding before analysis could protect patient identity. Further privacy-protecting coding methods that allow running prediction algorithms on coded data during protecting the identity of the patients are important for managing healthcare efficiently. There is a requirement for analysis and processing in a different manner by different methods. Performing resources exhausting operation while protecting privacy in a resource-constrained situation. Over and above healthcare are widely used by people with new privacy laws is a requirement of the hour.

7.2 Prediction can cause discrimination

Big data is usually employed in predictive information about other people's history. The information collected form of big data is increasingly becoming more advanced and it is becoming a tool for discrimination against people in different ways. A survey conducted shows that the information from Facebook and other web portals shows that information collected in this way discriminate people depending on rays, alcohol consumption, gender, etc, It causes concern discriminating people in this way by companies, educational institutions and other organizations may use this big data to discriminate people on human-oriented parameters.

7.3 Product sales

One of the major uses of big data is in marketing where marketing organizations place their products and services in targeting customers but when the customer is placed in a separate category based on their behavior there is a possibility of misuse. Despite the difficulties, marketing organizations depend on big data to target people through social media platforms like search engines and e-mails, and other applications. Sometimes due to personal hobbies, people tend to express their wishes and programs through e-mail and other social media platforms are causing some kind of anxiety with the user. For example, if we buy any products through online platforms they will target their customers through social media to sell their new products to them.

8. Existing ways to protect data from breach

8.1 Password and controlled access

A password is a form of digital access controlled security format to log in to e-mail, Facebook, and digital transactions like ATM and other online transactions to facilitate doing business and other social media platforms. The survey shows the history of the person's identity can be hacked and misused to their hacker's advantage. In the digital world, passwords consist of typographical characters used to authorize to approve the person to the computer system and other digital devices. While passwords can improve data security there are limitations for it is used. Passwords can be transferred from person to person without the permission of the data owners.

People with incomplete knowledge of computers and other electronic devices use very simple passwords to gain access to the electronic devices. This helps the hacker to break into their passwords easily and have

access to their computer system. More secure solutions to reset passwords to blocked portals or usually considered to be more tedious and it is difficult to be adopted. A solution for this is to adopt multi-factor authentication systems or passwords to log in to electronic devices is generally considered to be a more advanced and more secure method. Authentication historically requires a person to submit a password and a physical thing like credit card and debit card and more advanced systems use biometrics like fingerprinting or iris scanning as a means of authentication to provide access to it is the customer.

8.2 Data leakage prevention technology

Data prevention technology is a relatively new system of protecting the data of the customers or organization from people with malicious intent. Data protection involves the inspection of data files and their location classification and restriction of data movement from their within and outgoing internal networks by enforcing rules that are based on their data locations and classification of the files. Data protection technology is too stringent and makes end-users more frustrating for the staff concerned. Thus data protection technology is not so attractive to be made complain.

9. Conclusion

Big data has become a priced position for organizations like banking, educational institution, etc. This big data are being used by companies for analyzing customers' preferences and interest. It can be used for personal benefits and to launch misinformation about their customers. The main reason for this is the accessibility of big data in the form of digital format. So this is causing mistrust among customers about the organization they deal with.

REFERENCES

[1]. Boyd, Danah, and Kate Crawford, "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.”Information, Communication, and Society.

[2]. “Big Data is the Future of Healthcare”, Cognizant 20-20 insights, September 2012.

[3]. P.Kamakshi, “Survey on Big Data and its Privacy Issues”.

[4]. Duygusinanc, “A Survey on Security and Privacy Issues in Big Data”December 2015.

[5]. T. Vijey, A. Aiiad, "Big Data Security Issues Based on Quantum Cryptography and Privacy with Authentication for Mobile Data Center”, Procedia Computer Science, vol. 50, 2015.

[6]. “Big Data Analytics” Ericsson white paper,284 23-2311, August 2013.

[7]. Thomas M.Lenard and paulH.Dublin, “The Big Data Revolution: Privacy Considerations”, December 2013.

[8]. Big Data Analytics for Security Intelligence, September 2013.

[9]. Agarwal R., Srikant R., "Privacy-Preserving Data Mining.," in the proceedings of ACM SIGMOD conference.2000.

[10]. “Big Data Analytics” Ericsson white paper,284 23-3211, August 2013.

[11].Advantech (2013) Enhancing Big Data Security Retrieved from http://www.advantech.com.

[12]. IDC (2012). Big Data in 2020. Retrieved from http://www.emc.com/leadership/digital-universal/2020view/big-data02020.htm.

[13]. Sharma, S. Rise of Big Data and related issues. In Proceedings of the 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 17–20 December 2015; pp. 1–6.

[15]. Eynon, R. The rise of Big Data: What does it mean for education, technology, and media research? Learn. Media Technol. 2013, 38, 237–240. [CrossRef]

[16]. Wang, H.; Jiang, X.; Kambourakis, G. Special issue on Security, Privacy and Trust in network-based Big Data. Inf. Sci. Int. J. 2015, 318, 48–50. [CrossRef]

[17]. Thuraisingham, B. Big data security, and privacy. In Proceedings of the 5th ACM Conference on Data and Application Security and Privacy, San Antonio, TX, USA, 2–4 March 2015; pp. 279–280.

[18]. Rijmenam, V. Think Bigger: Developing a Successful Big Data Strategy for Your Business; Amacom: New York, NY, USA, 2014. 9. Big Data Working Group; Cloud Security Alliance (CSA). April 2013. Available online: https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_ Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf (accessed on 9 December 2015).

[19]. Dona Sarkar, Asoke Nath, "Big Data – A Pilot Study on Scope and Challenges", International Journal of Advance

[20].Big Data Security Issues and Challenges (PDF Download Available). Available from: https://www.researchgate.net/publication/275772328_Big_Data_Security_Issues_and_Challenges [accessed Apr 12 2018].