

A COMPREHENSIVE STUDY ON COGNITIVE DATA STORAGE USING BIG DATA APPROACH

Viola Disha Pais¹, Nishanka Seetharama² & Prof. Santosh Rebello³

Abstract: If your storage could analyse the real time data, the images, human sentiments and relationship how great would it. To do so the computer need to continuously analyse and learn, reason it out and make better decision just like how a human brain would do. IBM has introduced the concept of cognitive storage that is intended for constant examination and upgraded for unstructured information. The computers can be taught to learn the difference between high value and low value data. The main concept is to make computer learn what to remember and what to forget. It can consequently oversee information and conform to various necessities, extricate important data from information all by itself, adjust application prerequisites and match the cost of capacity to the estimation of the information so your business accomplishes huge investment funds.
Keywords: Cognitive Storage, Big Data, Cognitive Computing, Multitier Storage Unit, Neural networks.

1. INTRODUCTION

The idea of big data has been integrated in most domains. The digital learning which contributes mainly to emphasize on learning and teaching consists of data and analytics. The obstacle in handling such data can be a expensive. IBM introduced the "Cognitive Storage" technology to ensure that the most relevant information is always accessible. This technology determines the data, stores it in definite media, application of different levels of data security, schedules for the processing and reserves the different classes of data. This technology can be very helpful for online learning and is also very beneficial in the storage of relevant knowledge.

2. COGNITIVE COMPUTING

Cognitive Computing is based on self-learning systems which use machine-learning techniques to perform certain tasks in a high level way. It is the disguise of human thought processes in a computerized model that involves self-learning systems using data mining, pattern recognition and language processing to imitate the way the human brain works. The creation of automated IT systems that have problem solving capabilities without the help of humans is the goal of cognitive computing. These systems use machine learning algorithms and continually acquire knowledge from the data input in them by data mining for information. The systems are capable of predicting new problems and diagnosing feasible solutions to remove undesirable elements in patterns and in the processing of data. Cognitive computing is used in various Artificial Intelligence applications, which include neural networks, robotics and virtual reality. Cognitive Computing is anticipated in the transformation of lives, working and thinking and is therefore a significant tool.

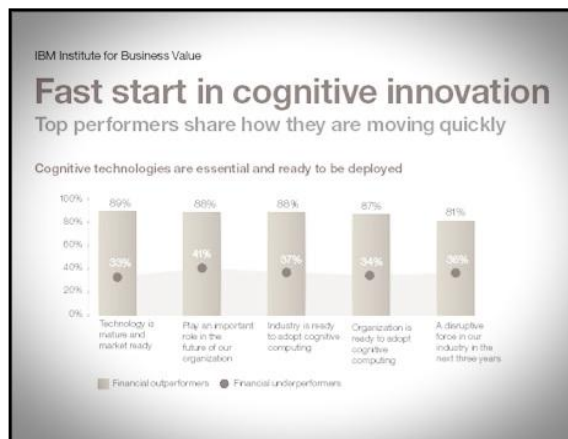


Figure 1. Cognitive Innovation

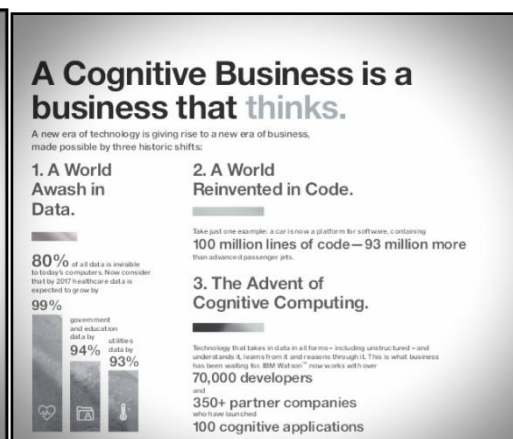


Figure 2. Cognitive Computing in Business

¹ Department of MCA, AIMIT, St. Aloysius College, Mangalore, Karnataka, India.

² Department of MCA, AIMIT, St. Aloysius College, Mangalore, Karnataka, India.

³ Department of IT and Bio-Informatics, AIMIT, St. Aloysius College, Mangalore, Karnataka, India.

3. BIG DATA

Big Data is perceived as massive data volume, used for social media sentiment analysis but in reality big data is about variety, velocity, validity and volume. Big data is about viewing all types of data sources as a business hurdle. It's not about having better insights and more data but it is more important to the right data than having the most data. The analytics on this is about making better decisions by having better queries, it is not just predictive analytics but to establish the benefit to be gained from the basics.[1]

Big Data is tremendously growing, not only in competence but also in complexity. As a result, it is used to enable forecasts to be made based on fresh data rather than relying on the old information. We need to explore new forms of data appropriation method to manage and integrate in order to make apt decisions. However, making a reasonable picture from infinite bits of information is not easy.[8] To maximize big data's value and to make it effective the focus needs to be on relevant data, with control on its quality to deliver the right decisions.

The reality the big data of today may not be the big data of tomorrow; what is entitled as 'big data' today will change priority as the tools and techniques to handle it evolve, the storage capacity increases and processing power improves.[10] Big data helps the management to stop looking at the past reactions and make predictive decisions about the future of your organization depending on relevant data.

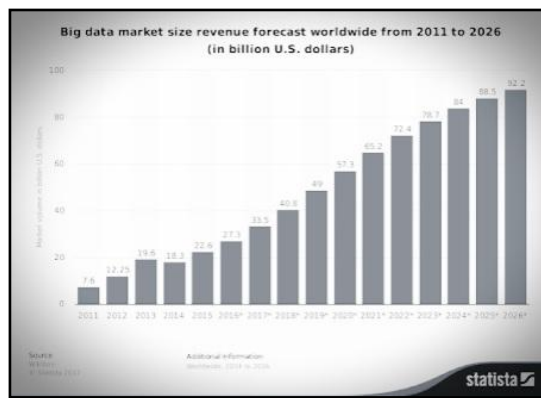


Figure 3. Statistics showing Big Data market revenue

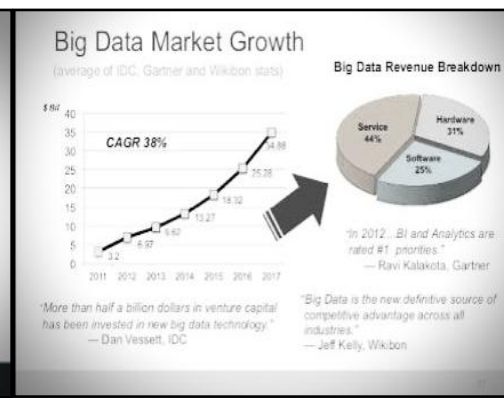


Figure 4. Big Data Market Growth

4. COGNITIVE DATA STORAGE USING BIG DATA

The cognitive storage considers a certain level of quality other than the regular approaches. Here the value of data is of grave importance for the basis of storing the grouping and management, placement and protection of data and the life cycle management of the data.[2] Cognitive storage is systematic because it stores the most relevant data with high redundancy and the lesser relevant data with less redundancy. Different properties of data are responsible for storing the datasets such as the value of data, data popularity, data life cycle and storage cost. The data placement and protection requirements are affected when emphasis is given on the value of the data and popularity.[3] Integrating the storage with computing power, determining the value of data, selecting data accessibility and security, moving data, evaluating data retrieval, updating selection criteria managing storage capacity are the various units involved in the enhancement of storage of Big Data.

In this system, it is important that the resources must pre-process the input data. The metadata of the data consists of the information about its relevance, popularity storage tier and repetitions. After the data is classified, it is sorted accordingly to determine the security level and the tier for each dataset in the Multitier Storage Unit. The data is selected by the selector and a data protection level is assigned to it for the initial placement in the appropriate storage tier.[6] Then the MTSU which consists of the migratory and storage capacity manager are responsible for taking into account the access pattern evaluation. The migratory is responsible to move the data according to the relevance allotted and also to maintain the protection level by modifying the redundancy schemes. The prime characteristic of Cognitive storage is to apply analytics engine for the continuous assessment of the value of data and predict the variations on which the classification process dependant. The storage manager assesses the redundancy of data and recommends extensions of the storage systems and deletes the lesser relevant data.



Figure 5. Cognitive Computing benefits based



Figure 6. Data Analytics in business strategies on life science executives.

5. STRATEGIC ADVANTAGE

The advantage will be that the system will decide independently on which storage medium is the most effective for a particular piece of data and will therefore select the cheapest which is fully compliant and fit for purpose – so where there is limited storage media, a low-level piece of correspondence can be kept on an inexpensive tape backup rather than a costlier piece of high tech. Cognitive storage considers distinctive benchmarks other than customary methodologies. In this kind the estimation of information has been given an extraordinary significance with the end goal of capacity arrangement and administration, position and assurance of information and the life cycle administration of the information. Cognitive storage makes the framework more proficient by putting away the most significant information with high excess and the less important information with less repetition. Different properties of information are in charge of putting away the datasets, for example, the estimation of information, information ubiquity, information life cycle and capacity cost. In Cognitive technology different properties of information are in charge of putting away the datasets, for example, the estimation of information, information ubiquity, information life cycle and capacity cost. This move to structures that can reason and learn is especially pertinent to the fundamental issue. According to IBM post “The memories you are recalling were captured because your brain automatically puts a high value on significant experiences, such as a beautiful sunset or an amazing dinner,” Simultaneously, your brain also automatically puts a low value or forgets irrelevant things like waiting at a traffic light or checking in for your flight. With cognitive storage, computers can do the same.

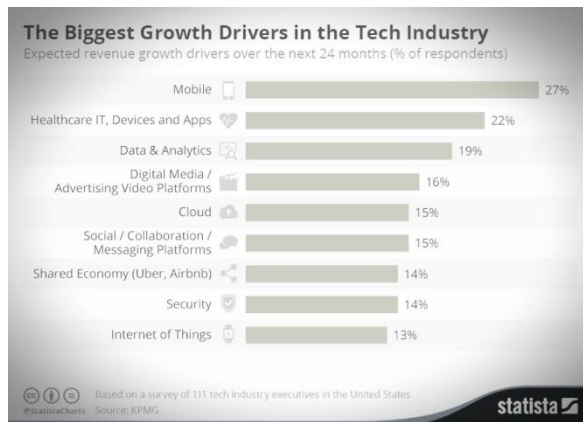


Figure 7. Growth in Tech Industry in Data and Analytics[11]

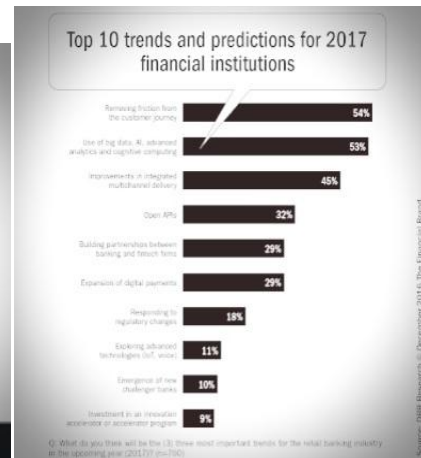


Figure 8. Trends in Financial Institutions

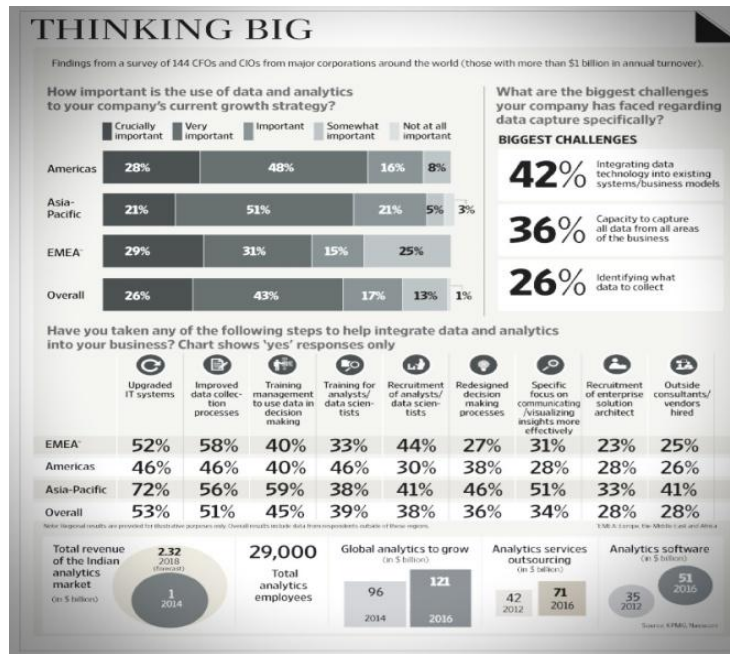


Figure 9. Data and Analytics in Company's Growth [14]

6. LIMITATION

Initially Cognitive systems were built to detect huge information for which it requires extensive measure of raw material. Vast worldwide associations may have an adequate pattern of substance and value-based and computerized data to get an incentive from machine learning calculations, however little and moderate size ones may battle to achieve minimum amount for the further developed frameworks. And one more crucial problem is that test for usage includes "preparing" the cognitive system. Training the cognitive system is a time consuming process and requires skilled resources to do the job. And the biggest challenge of all is cost to implement the system.

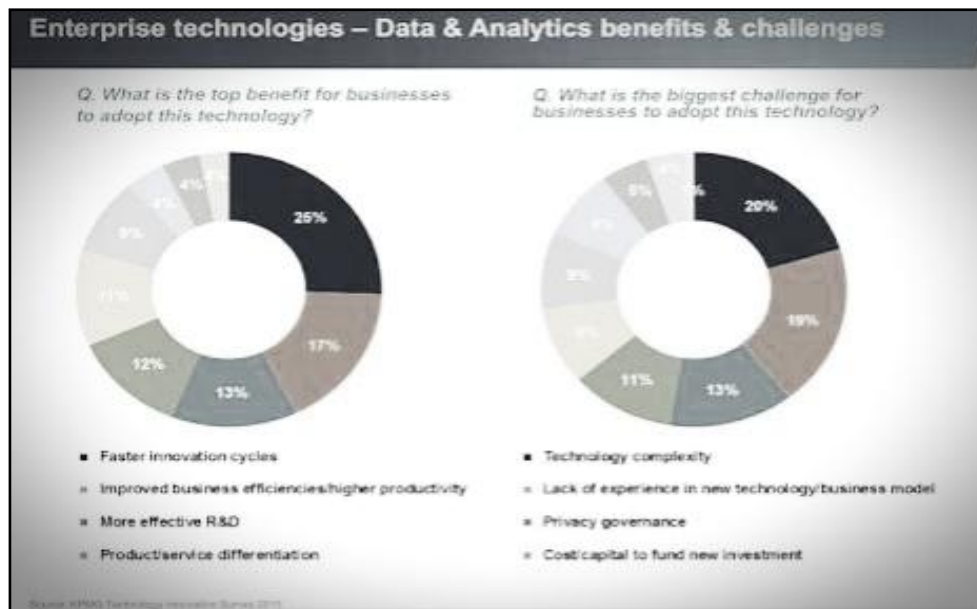


Figure 10. Data and Analytics challenges [13]

7. CONCLUSION

Machines may not have reached the point where they can make full decisions without humans, but they have certainly progressed to the point where they can make educated, accurate recommendations to us so that we have an easier time making decisions. Current machine learning systems have delivered tremendous benefits by automating tabulation and harnessing computational processing and programming to improve both enterprise productivity and personal productivity.

Cognitive systems will learn and interact to provide expert assistance to scientists, engineers, lawyers, and other professionals in a fraction of the time it now takes. While they will likely never replace human thinking, cognitive systems will extend our cognition and free us to think more creatively and effectively, and be better problem solvers.

8. REFERENCES

- [1] Make sense of big data .Learn its true value by making real-time analytics a reality. <https://www-03.ibm.com/systems/storage/big-data-storage/architect-your-future.html>.
- [2] Machine Learning, Cognitive Computing, And The Storage Industry. <https://thesanguy.com/2017/07/13/machine-learning-cognitive-computing-and-the-storage-industry>.
- [3] Reynolds H., "Big Data and Cognitive Computing" April 2016.
- [4] X.ci and X. Meng, "Big Data Management Concepts, Techniques and Challenges", Journal of Computer Research and Development, 2013.
- [5] Kaiser S. et al, "Big Data: Issues and Challenges", IEEE, 2012.
- [6] Cho S., "Fast Memory and Storage Architecture for the Big Data", IEEE, 2015.
- [7] Suthaharan S., "Network Intrusion Prediction with Machine Learning", Performance Evaluation Review, Vol.41, No.4, March 2014.
- [8] Zadrozny P. et al, "Big Data Analytics using Splunk, Apress, 2013.
- [9] Mininath R. Bendre, Vijaya R. Thool, "Analytics challenges and applications in big data environment: a survey", Journal of Management Analytics, pp. 1, 2016, ISSN 2327-0012.
- [10] Cherubini G, Jelitto J., Venkatesan V. , "Cognitive Storage for Big Data", IEEE Computer 49(4), 43-51, 2016
- [11] Harris R., "Smart Storage for Big data", April 2016.
- [12] Clapperton G., "IBM To Slash cost of Big data with "Cognitive Storage", April 2016.