

Future of Big Data

Ms. Priya Chaudhari¹, Ms. Binita Patel²

¹Final year Student B.E., Department of Computer engineering, SVIT Vasad, Gujarat, India

²Final year Student B.E., Department of Computer engineering, SVIT Vasad, Gujarat, India

Abstract – As quoted by CNBC – “Data is the new oil”. Big data has become a business issue or at least an issue that business people are aware of. Due to the advancement of new technologies, devices, social media, the amount of data produced is growing rapidly every year. For instance, The Large Hadron Collider, when operational last year, generated 1GB of data per second. The amount of data produced by mankind from the beginning of time till 2003 was 5 billion gigabytes. The same amount was created in every two days in 2011, and in every ten minutes in 2013. This rate is still growing enormously. Big data is a collection of large datasets that requires advanced tools, techniques and frameworks rather traditional computing techniques. Big data is evolving rampantly at science symposiums, playing a most vital role across industries and business firms. As a result the data are mined for deeper insights and improve business functions, sales, customers review and satisfaction, privacy and security.

Key Words: Big Data, Hadoop, Cyber security, Machine learning, IoT, Cloud, Spark

1. INTRODUCTION

Almost everyone has agreed that a big data has taken the business world by a storm, but what’s next? Will data continue to grow? What technologies will develop around it?

Big data has provided firms and institutions with better sources for analysis, marketing, business decisions, as well increasing the ability to see those pressure points and correct issues before customers are even aware of them. Although, with more data points being collected from customers, and employees, more and more firms are facing the challenge of how data can be retrieved and analyzed faster and more efficiently. The concept of big data fabric helps firms to address these problems in a way to store, analyze and stream data that will be efficient to retrieve actionable information and leverage the value of the data that we get from customers. Data will continue to grow, considering that the number of hand handled devices and Internet-connected devices will grow exponentially.

1.1 Challenges

- Store, handle, analyze
- Increasing volume
- Security

- Integrating collections of data to yield actionable information
- Employing advanced tools to manage and retrieve data efficiently
- Staffing shortages
- Forecasting trends

2. Trends and Predictions of Big Data

It’s clear that the big data market will grow but how organizations will be using this big data is not that clear. Along with current technologies new technologies will be surfaced to improve the big data market. Below are some trends that are likely to shape the big data market in future:

1. Open source
2. Machine Learning
3. Edge Analytics & Computing
4. Predictive Analytics
5. Cloud Computing
6. Intelligent Security
7. Job trends

2.1. Open Source

Open source applications like Apache Hadoop, Spark, R programming, Map reduce and more are dominating the big data space. Hadoop has quickly become synonymous of big data, but it is difficult to master hadoop, therefore lots of projects have failed; but now there is Spark, and Hadoop has come quite a long way. Hadoop is becoming affordable general purpose solution. But spark, a new framework is almost 100 times faster than hadoop. Whereas, R is a programming language and software environment for statistical analysis, graphical representation and reporting.



Fig-1: Open sources of Big Data

platforms like hadoop whereas others are dependent on vendor products that include big data analytic properties. As reported Cybersecurity Ventures, cyber security expenditure will exceed \$1 trillion from 2017 to 2021.

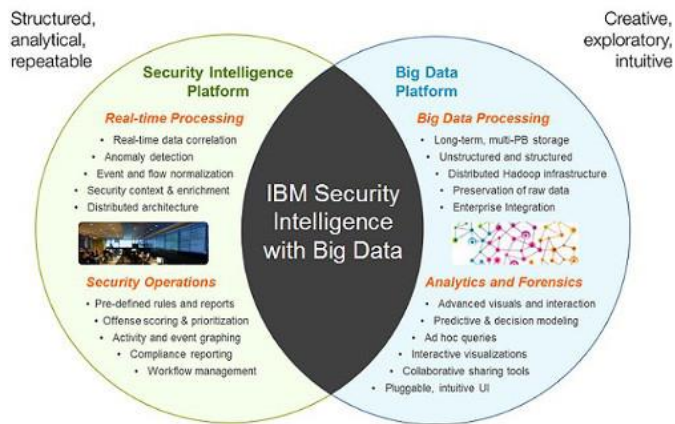


Fig-4: Intelligent Security

2.7. Job Trends

Due to increased demands there is need to bridge the talent gap in data analytics. Therefore, academic institutions are compiling degree programs in data science, boot camp style schools are providing training in marketable skills. According to an industry estimate data scientist will receive twice the remuneration received by a programmer. Also companies are more likely to appoint chief data officers (CDO).

DATA/DATABASE ADMINISTRATION (B)			
Big Data Engineer	\$129,500 - \$183,500	\$135,000 - \$196,000	5.8%
Database Manager	\$118,000 - \$170,500	\$122,250 - \$177,000	3.7%
Database Developer	\$103,250 - \$153,250	\$108,000 - \$161,500	5.1%
Database Administrator	\$ 95,750 - \$142,750	\$ 98,500 - \$148,500	3.6%
Data Analyst/Report Writer	\$ 74,500 - \$114,500	\$ 77,500 - \$118,750	3.8%
Data Architect	\$127,250 - \$175,500	\$131,250 - \$184,000	4.1%
Data Modeler	\$106,750 - \$155,500	\$111,000 - \$161,500	3.9%
Data Scientist	\$109,000 - \$153,750	\$116,000 - \$163,500	6.4%
Data Warehouse Manager	\$123,750 - \$172,000	\$129,000 - \$179,000	4.1%
Data Warehouse Analyst	\$105,000 - \$152,000	\$107,500 - \$155,750	2.4%
Business Intelligence Analyst	\$113,750 - \$164,000	\$118,000 - \$171,750	4.3%
Electronic Data Interchange (EDI) Specialist	\$ 76,500 - \$115,500	\$ 77,500 - \$118,750	2.2%
Portal Administrator	\$ 94,000 - \$134,250	\$ 96,000 - \$138,750	2.8%

Fig-5 Predicted Salaries

3. CONCLUSION

Gartner predicts that by 2020 big data and Internet of Things will be used to reinvent, digitize and eliminate 80% of business processes. Therefore, the impact that big data will have on various businesses cannot be overlooked. To truly uncover the power of big data, organizations must identify how to use their data to build reporting and analytics for

their operations. Studying all the latest leading big market analysis and research, it can be assumed that 15% of IT organizations will move to cloud based service platforms and service market is anticipated to rise about 35% by 2021. Summarizing, Peter Sondergaard, Senior Vice President of Gartner Research famously stated, — "Information is the oil of the 21st century and analytics is the combustion engine" ..

REFERENCES

- [1] Huawei Noah's Ark Lab, Hong Kong Science Park, Shatin, Hong Kong "Mining big data: current status and forecast To the future" ACM SIGKDD Explorations Newsletter Volume 14 Issue 2, December 2012.
- [2] "BIG DATA MINING" VIT University, Vellore Yash Gupta, Mustafa Bohra, Mastan Patel, Jeet Kumar IRJET publications Volume: 03 Issue: 11 | Nov -2016
- [3] Lizhe Gregor, von Laszewski, Andrew YoungeXi, HeMarcel KunzeJie, TaoCheng Fu "Cloud Computing: a Perspective Study" Article First Online: 03 June 2010 DOI: 10.1007/s00354-008-0081-5
- [4] Ken Birman, Gregory Chockler, Robbert van Renesse "Toward a cloud computing research agenda" Newsletter ACM SIGACT News archive Volume 40 Issue 2, June 2009
- [5] Amir Gandomi, Murtaza Haider "Beyond the hype: Big data concepts, methods, and analytics" Volume 35, Issue 2, April 2015, Pages 137-144
- [6] Stephen Kaisler, Frank Armour, J. Alberto Espinosa, William Money "Big Data: Issues and Challenges Moving Forward"
- [7] Data Mining: Practical Machine Learning Tools and Techniques, Second Edition By Ian H. Witten, Eibe Frank
- [8] Varian, Hal R. "Big Data: New Tricks for Econometrics" The Journal of Economic Perspectives, Volume 28, Number 2, Spring 2014, pp. 3-27(25) in American Economic Association