



## Big Data Analytics and Internet of Things (IoT): a Review

---

Ashok Deokar

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

November 26, 2021

# Research Paper

On

## “Big Data Analytics and Internet of Things (IoT): A Review”

Prof. Ashok Deokar

Asst. Professor

Dr. D Y Patil School of MCA

Savitribai Phule Pune University.

a\_deokar@rediffmail.com

### Abstract

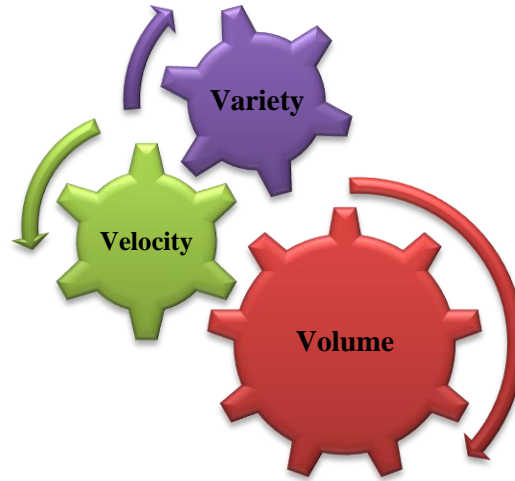
In today's techno savvy world everything is digitalized. Everything is on the tips of people. Trends are changing in seconds. The different trend in the current modern world is that all the devices used by common people are connected with the internet and the reason of this is to share the information or data without any human communication and improving the excellence of our daily lives. The purpose of this research was to provide understanding about big data analytics and Internet of Things (IoT).

**Keywords: Big Data, Internet of Things (IoT), Artificial intelligence (AI)**

### Introduction

Big data analytics is a buzz word in today's modern era. This is used for data sets having huge size or type and which is very far from the capability of traditional relational databases to gather, manage and process the data with low potential. “The term Big Data means, large volume of data that are generated by devices, sensors, social media, health care applications, various other software applications and digital devices.” [12] “Big Data has recently become more prominent in the IT technology, where it helps in product optimization, improves decision making and saves energy.” [10] Artificial intelligence (AI), mobile and the Internet of Things (IoT) are managing data complexity through new forms and sources of data. Big data generated from sensors, devices, video/audio, networks, log files, transactional applications, web, and social media much of it produced in real time and at a very huge scale.

Big Data is huge in volume, up till now increasing exponentially with period. Volume, Velocity and variety are the characteristics of big data. These three Vs of big data are shown in fig.1 and explained below:



**Fig. 1 Characteristics of Big Data**

- **Volume**

The amount or volume of data matters for better result and analysis. Huge amount of data is generated on daily basis from various sources like social media, business processes, sensors, mobiles etc. This can be facts and figures of unknown value, such as Twitter data feeds, click streams on a webpage or a mobile app, or sensor-enabled equipment. For some organizations, this might be tens of terabytes of data and for some organizations it may be hundreds of petabytes.

- **Velocity**

Velocity is nothing but the accumulation of the data on high speed. It is the fast rate at which data is received and performed. It contains the linking of incoming data sets speeds, rate of change and activity bursts. As the main aspect of big data is to provide data rapidly, velocity plays a vital role.

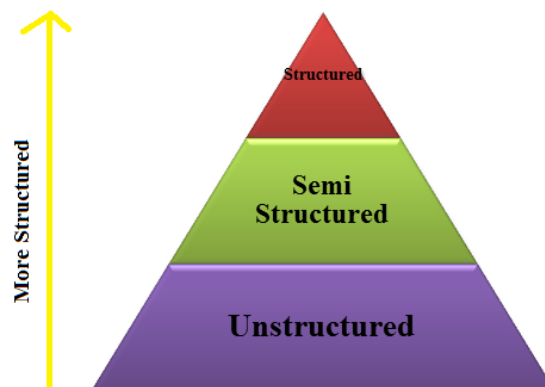
- **Variety**

Variety refers to the data that are available different sources and in different structures. It can be structured, semi-structured and unstructured. Data can be in the form of emails, photos, text, audio, and video, monitoring devices, PDFs etc.

## ❖ Types of Big Data

There are three types of big data structures mentioned as follows:

1. Structured
2. Unstructured
3. Semi-structured



**Fig. 2 Types of Big Data: Data Structures**

1. **Structured:** The data which can be stored, processed and accessed in form of structure schema is structured data. It follows a consistent order and it is designed in such a way that it can be easily retrieved and used by a person or a computer. As data of relational database management system is stored in tabular format (Which is a structured one) can called as structured data.
2. **Semi-structured:** Semi-Structured Data can be considered as another form of Structured Data. It inherits a few properties of Structured Data, but the major part of this kind of data fails to have a definite structure. It is a form of structured data that does not follow with the proper structure of data models related with relational databases or other forms of data tables, but even so contain tags or other indicators to distinct semantic features and enforce hierarchies of records and fields within the data.
3. **Unstructured:**

Unstructured data is data that either does not have a predefined data model or is not organised in a pre-defined manner. Unstructured data is usually text-heavy, but may comprise data such as dates, numbers, and facts also. These results in anomalies and uncertainties that make it challenging to comprehend using traditional programs compared to data stored in structured databases.

“Predictive analytics, which deals mostly with structured data, overshadows other forms of analytics applied to unstructured data, which constitutes 95% of big data.”  
[9]

### **Advantages of Big Data Analytics**

Organizations practice big data analytics systems and software to take data-driven decisions that can improve and increase profits. The advantages may comprise more effective marketing with new strategies, new revenue opportunities, customer personalization and improved operational efficiency. Some more advantages are listed below:

- Cost Optimization
- Improve efficiency
- Foster competitive
- Boost sales and retain customer loyalty
- Innovate
- Focus on the local environment
- Control and monitor online reputation

### **Internet of Things (IoT):**

Internet of Things (IoT) denotes to a system of connected physical items with the help of internet. “It is a web-based network, which connects smart devices for communication, data transfer, monetary exchange, and decision-making. Both the number of communication channels and the volume of data transmitted are increasing exponentially along with the number of devices that are connected to this network.” [11] The ‘thing’ in IoT can denote to an individual or any device which is allocated with the help of IP address. A ‘thing’ gathers and transmits data over the internet deprived of any physical involvement with the help of embedded technology. It supports them to communicate with the peripheral environment or

internal states to take the decisions. “IoT has created unprecedented opportunities that can help increase revenue, reduce costs, and ameliorate efficiencies, collecting a huge amount of data alone is insufficient.” [6] “IoT technologies have been incorporated into various important domains in our life. Over the past years, many traditional domains such as manufacture industry, healthcare or energy have become IoT-based and gained the capability of communicating among machines and human, as well as production of enriched data.”[4]

### **Role of Big Data in Internet of Things (IoT):**

When organizations are taking hold of the data for analysis purpose, IoT is performing a major source for that data, and for that reason the role of big data in IoT comes into the picture. Big data analytics is developing as an important to analysing IoT generated data from “connected devices” which helps to take the initiative to improve decision making. “IoT interacts with big data when voluminous amounts of data are needed to be processed, transformed, and analysed in high frequency.” [6] “The Internet of Things is generating an enormous amount of data. Analysing and managing that data requires programming and statistical approaches. Big Data technology operates on this massive data and pushes new products, applications, future research and developments to improve decision making.” [10] “Today more and more technology is getting evolved, research have now proven how machines can communicate with human.” [1] The role of big data in IoT is to process a bulky data on an actual basis and storing them with different types of storage technologies. IoT big data processing follows four phases:

#### **Phase 1:**

A huge amount of unstructured data is produced by IoT devices which are gathered in the big data system. This IoT engendered big data mostly depends on big data’s three V’s factors which is volume, velocity, and variety.

#### **Phase 2:**

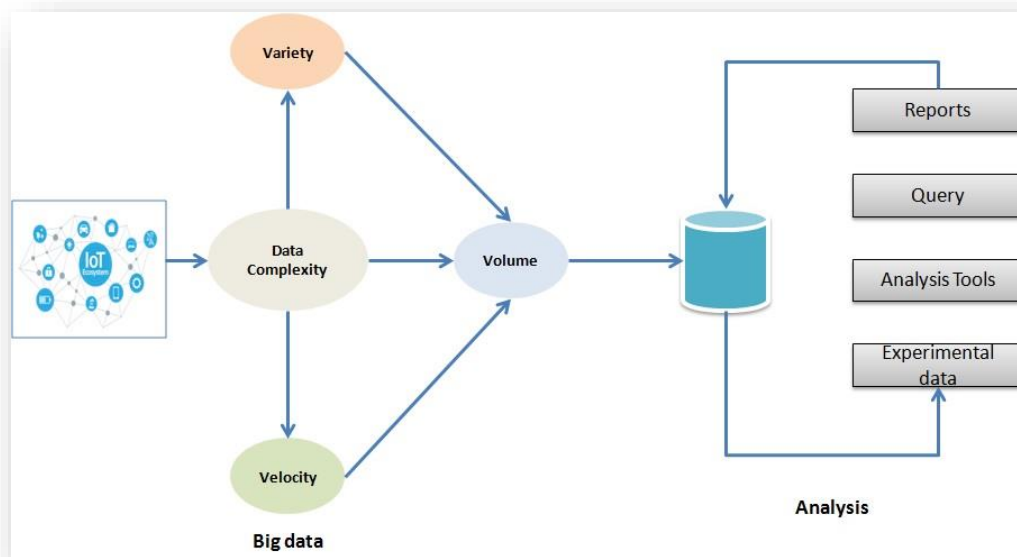
In the big data system which is mostly a shared distributed database, the massive amount of data is stored in big data files.

### Phase 3:

Investigating or analysing the stored IoT big data by using analytical tools.

### Phase 4:

After analysing the data generating the reports.



**Fig 3: IOT Big Data Processing** [Source: Google Images]

Subsequently in IoT the unstructured data are gathered with the help of the internet, therefore, big data for the internet of things require fast analysis with large queries to increase quick visions from data to make fast decisions. Hence the need for big data in IoT is persuasive. “Voluminous amounts of data have been produced, since the past decade as the miniaturization of Internet of things (IoT) devices increases. However, such data are not useful without analytic power. Numerous big data, IoT, and analytics solutions have enabled people to obtain valuable insight into large data generated by IoT devices.” [2] IoT and big data both the technologies transmit inter-dependency and need further development. “There are a few areas of concern and security and privacy and data collection efficiency are probably the most difficult problems are facing by the users.” [5]

## Conclusion:

In today's world dealing with real-time data is in great demand and it is also a necessity. Remarkably, both Big data Analytics and IoT are in demand. The researcher concluded that the collective supremacy of IoT and Big data Analytics can provide innovative starts and applications in every business segment. In addition to that, it has an adequate amount of prospective to bring innovatory vicissitudes in many aspects of social lives. Big data Analytics and IoT are two distinct concepts, but they depend on each other for final achievement. Both give emphasis to the need for transforming data into concrete understandings that can be performed.

## Bibliography

1. L. Kaul and R. H. Goudar, "Internet of things and Big Data - challenges," *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, Coimbatore, 2016, pp. 1-5, doi: 10.1109/GET.2016.7916854.
2. M. Marjaniet *al.*, "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges," in *IEEE Access*, vol. 5, pp. 5247-5261, 2017, doi: 10.1109/ACCESS.2017.2689040.
3. M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2923-2960, Fourthquarter 2018, doi: 10.1109/COMST.2018.2844341.
4. Ge, Mouzhi & Bangui, Hind & Buhnova, Barbora. (2018). Big Data for Internet of Things: A Survey. *Future Generation Computer Systems*. 87. 10.1016/j.future.2018.04.053.
5. K. R. Kundhavai and S. Sridevi. (2016). IoT and Big Data-The Current and Future Technologies: A Review. *IJCSMC*, Vol. 5, Issue. 1, January 2016, pg.10–14
6. Ahmed, Ejaz & Yaqoob, Ibrar & Hashem, Ibrahim & Khan, Imran & Ahmed, Abdelmutilib Ibrahim Abdalla & Imran, Muhammad & Vasilakos, Athanasios. (2017). The role of big data analytics in Internet of Things. *Computer Networks*. 129. 10.1016/j.comnet.2017.06.013.



7. Habib urRehman, Muhammad & Yaqoob, Ibrar & Salah, Khaled & Imran, Muhammad & Jayaraman, Prem Prakash & Perera, Charith. (2019). The Role of Big Data Analytics in Industrial Internet of Things. *Future Generation Computer Systems*. 10.1016/j.future.2019.04.020.
8. Somayya Madakam, R Ramaswamy and Siddharth Tripathi, "Internet of Things (IoT): A literature review", *Journal of Computer and Communications*, vol. 3, no. 05, pp. 164, 01 2015.
9. Amir Gandomi and MurtazaHaider, *Beyond the hype: Big data concepts methods and analytics*, vol. 35, 2015.
10. S. Khare and M. Totaro, "Big Data in IoT," *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kanpur, India, 2019, pp. 1-7, doi: 10.1109/ICCCNT45670.2019.8944495.
11. Fayeem Aziz, Stephan K. Chalup and James Juniper. Big Data in IoT Systems. Interdisciplinary Machine Learning Research Group The University of Newcastle
12. S. ArulanandaJothi, J. Abdul Samath, K. Anandharaj (2019). A Study On IoT And Data Analytics In Healthcare Systems. *International Journal Of Scientific & Technology Research* Volume 8, Issue 10, October 2019