

# Swift Towards Big Data Analytics

S . Suneetha<sup>1</sup>

Research Scholar

Department of Computer Science

SPMVV, Tirupati

Dr . M . Usha Rani<sup>2</sup>

Professor

Department of Computer Science

SPMVV, Tirupati

**Abstract--**The promise of data-driven decision making is recognized broadly in this contemporary world and we've immigrated to the New Era of 'Big Data'. 'Big Data' are large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable capture, storage, distribution, management and analysis of information. The Analytics associated with Big Data is described by 5 major parameters: Volume, Velocity and Variety, Veracity and Value. Volume is growing and so are the veracity and other issues. Big Data Analytics drives nearly every aspect of our modern society including financial trading, transportation, manufacturing, mobile services, health care, revolutionizing the scientific research in Data Mining especially. Heterogeneity, Scale, Timeliness, Privacy, Provenance and Visualization problems with Big Data impede progress at all stages of the Big Data Analytics pipeline right from data acquisition to result interpretation that can create value from data. Fundamental research in Big Data Analytics should be encouraged and supported towards addressing these technical challenges to achieve potential benefits of Big Data to the fullest.

**Keywords--**Big Data; Analytics; Veracity; Provenance; Visualization; Privacy; Heterogeneity; Timeliness; Scale

## I. FOREWORD

We are in the Era of Big Data now. In a broad range of application areas data is collected at an unprecedented scale and the paradigm is termed as, 'Big Data'. 'Big Data' is high volume, high velocity and/or high Variety information assets that require new forms of processing to enhanced decision making, insight discovery and process optimization. Analytics are the foundation for a differentiated customer experience and more meaningful engagement as they can help to gain insights into the data. 'Big Data Analytics' is the application of advanced analytic techniques to very large, growing, diverse datasets with multiple autonomous sources to discover patterns and other useful information for informed decision making.

The misnomer is in the 'bigness' of data and the emergent discussions focus on how 'smart' the data is, rather than how 'big' it is. Big Data is a wrapper of different types of granular data. With the fast development of networking, data storage, collection capacity and computation, Big Data are now rapidly expanding in all application domains including physical, biological and biomedical sciences.

Big Data is generated from an increasing plurality of sources, including Internet Clicks, mobile transactions, user-generated content and social media as well as purposefully generated content through business transactions or sensor networks,. In addition, genomics, health care, engineering, operations, management, the industrial internet and finance add to big data pervasiveness. These data require the use of powerful computational techniques to unveil trends and patterns within and between these extremely large socioeconomic data sets. Now, insights gleaned from such data-value extraction can meaningfully complement static ones, adding depth and insight from collective experiences and doing so in real time narrows both information and time gaps.

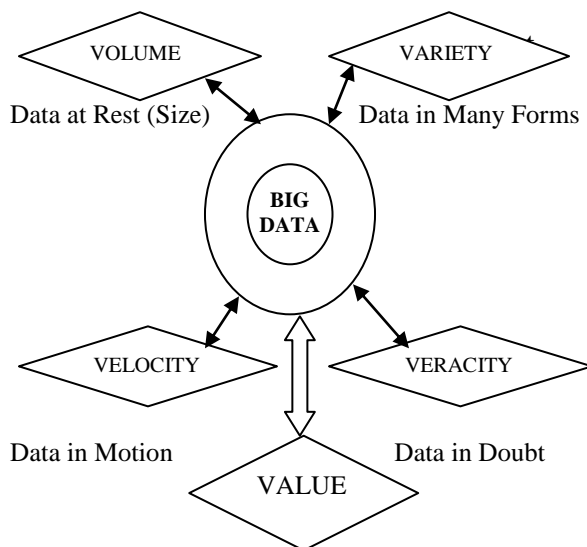
The promise and the goal of strong management research built on Big Data should be not only to identify correlations and establish plausible casualty but to reach Consilience i.e., convergence of evidence from multiple, independent and unrelated sources, leading to strong conclusions. Big Data offers exciting new prospects for such consilience due to its unprecedented volume (of data over multiple periods), micro-level detail and multi-faceted richness. Big Data requires exceptional techniques to efficiently process large quantities of data within tolerable elapsed times.

The rest of this study paper is organized as follows: Section II highlights 5 major V's for designing a Big Data Strategy. Section III lists Applications where Big Data Analytics are applied the most. Section IV provides a brief overview of Phases of a Big Data Analytics Pipeline along with the general research challenges in each phase. Technical Challenges to overcome are presented in Section V followed by Executive Summary.

## II. MULTI V-MODEL OF BIG DATA

The term 'Big Data' is coined by Roger Magoulas from O'Reilly media in 2005. Research on Big Data emerged in 1970s but has seen an explosion of publications since 2008.

Fig. 1. 5V's of Big Data Strategy



The first definition of Big Data was developed by Meta Group (now part of Gartner) by describing its 3 characteristics called “3V”: Volume, Velocity, and Variety. Based on data quality, IBM added a 4<sup>th</sup>V called: Veracity. However, Oracle added another V called: Value, highlighting the added value of Big Data.

To understand the phenomenon of Big Data, it is described using the 5V's that follow:

◆ ‘Volume’ refers to the vast amounts of data generated every second. Zettabytes, Yottabytes or Brontobytes of data is too large to store and analyse using traditional database technology. With Big Data technology, these data sets can be stored and used with the help of distributed systems, where parts of the data is stored in different locations and brought together by software.

◆ ‘Velocity’ refers to the speed at which new data is generated and the speed at which data moves around. Social media messages go viral in seconds, credit card transactions are checked for fraudulent activities at high speed or it takes milliseconds for the trading systems to analyse social media networks, to pick up signals that trigger decisions to buy or sell shares. Big Data technology analyses the data while it is being generated, without even putting it into databases.

◆ ‘Variety’ refers to different types of data that is used. In the past, the focus was on structured data that nearly fits into tables or relational databases. In fact, 80% of the world's data now is unstructured and therefore cannot be easily put into tables (photos, videos or social media updates). With Big Data technology, different types of data (structured, semi structured & unstructured) can be harnessed including messages, social media conversations, sensor data, voice or video recordings and bring them together with more traditional structured data.

◆ ‘Veracity’ refers to the messiness or trustworthiness of the data i.e., the uncertainty associated with the data. Veracity

arises due to data uncertainty, process uncertainty and/or model uncertainty. With sheer volume of data in many forms, quality and accuracy are less controllable. But, Big Data and analytics technology allows working with these types of data.

◆ ‘Value’ is the most important V of Big Data as access to Big Data is useless unless it can be turned into Value. The defining parameter of Big Data is the fine grained nature of the data itself, thereby shifting the focus away from Volume to Value.

To develop a Big Data strategy, multi-V model should also include: Volatility, Variability and Visualization.

◆ Volatility is to be understood due to volume, variety, and velocity of Big Data. For some sources, the data will always be there; for others, this is not the case. Understanding what data is out there and for how long can help one to define retention requirements and policies for Big Data.

◆ ‘Variability’ is defined as the “variance in meaning, in lexicon”. It is often confused with variety and variability means that the meaning is changing rapidly. In the same tweets a word can have totally different meanings; the algorithms need to understand the context and should decipher the exact meaning of a word in that context. This is still very difficult. Variability is thus very relevant in performing sentiment analysis and Big Data is extremely variable.

◆ ‘Visualization’ is making vast amount of data comprehensible in a manner that is easy to understand and read. With the right analyses and visualizations, raw data can be put in to use otherwise raw data remains essentially useless. Visualizations are readable and understandable complex graphs that can include many variables of data rather than ordinary graphs or pie charts. Visualizing might not be the most technological difficult part; but it is hard and the most challenging part of Big Data. Telling a complex story in a graph is very difficult and extremely crucial also. There are Big Data startups that focus on this aspect and in the end, visualizations will make the difference. One of them is Ayasdi and they use topological data analysis to discover patterns in vast amounts of data via stunning 3D graphs. In the future this will be the direction to go, where visualizations help organizations answer questions they did not know to ask even.

Proving the Value by optimally managing Data Volume, Variety, Velocity, Veracity and offering insights with compelling Visualization make up the Mantra of the 6V's of Big Data.

### III. APPLICATIONS OF BIG DATA

Big Data is one of the megatrends that will impact every application domain and every one in one way or another. The following are the awesome ways Big Data is used to change the data-driven contemporary world today:

#### ● Understanding and Targeting Customers

Big Data are used to better understand customers, their behaviors and patterns.

Eg: - Government election campaigns can be optimized using Big Data analytics. It is believed that, Narendra Modi's win in 2014 parliamentary election campaign was due to his team's superior ability to use big analytics.

#### ● *Understanding and Optimizing Business processes*

Big Data are also increasingly used to optimize business processes. Retailers can optimize their stock based on predictions generated from social media data, web search trends and weather forecasts.

Big Data analytics are used in supply chain delivery route optimization. Geographic positioning & radio frequency identification sensors are used to track goods or delivery vehicles and optimize routes by integrating live traffic data etc., HR business processes are also being improved using Big Data analytics. This includes optimization of talent acquisition as well as the measurement of company culture and staff engagement using Big Data tools.

#### ● *Personal Qualification and Performance Optimization*

Individuals can even benefit from the Big Data generated from wearable devices like smart watches or smart bracelets.

E.g.: - Big Data tools and algorithms help to find the most appropriate matches.

Human capital analytics including workforce management and employee performance optimization are also hot uses of Big Data. It helps to put performance in context so that improvement measures make more sense and are more effective. For e.g., If a restaurant server has fewer days than peak performance and if only drop in that employee's performance is considered, punishment or additional training is the best corrective measure. However, with Big Data one can see that the weather was bad and thus there were fewer patrons in the restaurant on those days, or that there was a big event in another part of the city that suppressed overall sales, or the construction work in front of the restaurant slowed store traffic perhaps. In any of those circumstances, no corrective measures are needed to make the employee more productive.

#### ● *Manufacturing*

Huge improvements in supply planning and boost product quality are the greatest benefits of Big Data for manufacturing. Big Data provides an infrastructure for transparency in manufacturing industry, the ability to unravel uncertainties such as inconsistent component performance and availability. Predictive manufacturing as an applicable approach toward near-zero downtime and transparency requires vast amount of data and advanced prediction tools for a systematic process of data into useful information. A conceptual framework of predictive manufacturing begins with data acquisition where different type of sensory data is available to acquire such as acoustics, vibration, pressure, current, and voltage and controller data. Vast amount of sensory data in addition to historical data construct the Big Data in manufacturing. The generated Big Data acts as the input into predictive tools and

preventive strategies such as Prognostics and Health Management (PHM).

#### ● *Improving Scientific Research*

Scientific research has been revolutionized by Big Data. The opportunities are enormous, so are the challenges, created by entirely new applications and by the relentlessly increasing volume, velocity, and variety of data. "Big Data" puts computer science & research at the center of advances in every imaginable field.

#### ● *Education*

Big Data has the potential to revolutionize not only just research, but also education. A complex world demands skilled and knowledgeable citizens and educational institutions at every level struggle to graduate students who can meet these requirements. As costs rise and funding shrinks, educators and administrators need deeper insight into approaches that work. Analytics provides them with the tools to measure performance and ensure students acquire the skills to succeed. Business analytics solutions help primary, secondary and higher education professionals: to promote academic success, measure learning effectiveness, manage financial performance and reduce risk and complexity.

#### ● *Financial Trading*

Big Data finds a lot of use in High-Frequency Trading (HFT) today. Majority of equity trading takes place via Big Data algorithms that take into account signals from social media networks and news web sites to make buy & sell decisions in a split of seconds.

#### ● *Improving Sports Performance*

Most elite sports have now embraced Big Data analytics. E.g.: - IBM Slam Tracker tool for tennis tournaments. Video analytics are used to track the performance of every player in a football or baseball game, and sensor technology in sports equipment such as basket balls or golf clubs allow to get feedback (via smart phones and cloud servers) on the game so as to improve it. Many elite sports teams also track athletes outside of the sporting environment – smart technology is used to track nutrition and sleep, as well as social media conversations to monitor emotional wellbeing.

#### ● *Improving Health Care and Public Health*

The computing power of Big Data analytics enables decoding entire DNA strings in minutes to better understand & predict disease patterns and to find new cures as well. It helps to monitor and predict the developments of epidemics disease outbreaks.

The use of Big Data analytics can reduce the cost of health care while improving its quality by making care more preventive & personalized with more extensive home-based continuous monitoring. Integrating data from medical records

with social media analytics enable monitoring of epidemic outbreaks in real time.

#### ● Improving Security and Law Enforcement

Big Data is used heavily in improving security and enabling law enforcement. Big Data techniques are used to detect and prevent cyber attacks. Police forces use Big Data tools to catch criminals and predict criminal activity even. Credit card companies use Big Data to detect fraudulent transactions.

#### ● Improving and Optimizing cities and countries

Big Data is used to improve many aspects of cities and countries as well. It allows cities to optimize traffic flows based on real time traffic information as well as social media and weather data. A number of cities are currently piloting Big Data analytics with the aim of turning themselves into smart cities, where the transport infrastructure and utility processes are joined up.

E.g.: - A bus would wait for a delayed train and traffic signals predict traffic volumes and operate to minimize traffic jams.

#### ● Optimizing Machine and Device Performance

Big Data analytics help machines and devices to become smarter and autonomous.

E.g.: - Big Data tools are used to operate Google's self-driving car. They are also used to optimize energy grids using data from smart meters. They can even be used to optimize the performance of computers and data warehouses.

In a similar vein, persuasive cases that made for the value of Big Data for urban planning (through fusion of high-fidelity geographical data), intelligent transportation (through analysis and visualization of live and detailed road network data), environmental modeling (through sensor networks ubiquitously collecting data), energy saving (through unveiling patterns of use), smart materials (through the new materials genome initiative, computational social sciences (a new methodology fast growing in popularity because of the dramatically lowered cost of obtaining data), financial systemic risk analysis (through integrated analysis of a web of contracts to find dependencies between financial entities), homeland security (through analysis of social networks and financial transactions of possible terrorists), computer security (through analysis of logged information and other events, known as Security Information and Event Management (SIEM)), and so on.

Though the potential benefits for applications of Big Data are real and significant, and some initial successes have already been achieved there remain many technical challenges that must be addressed to achieve the promised benefits of Big Data.

## IV. BIG DATA ANALYTICS PIPELINE

The analysis of Big Data involves multiple distinct phases as shown in the figure below, each of which introduces some general challenges.

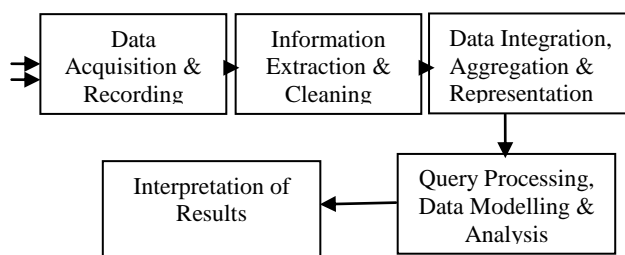


Fig. 2. Major Phases of Big Data Analytics

### A. Data Acquisition and Recording

Big Data is recorded from varied data sources. Much of this data will be of no interest, and it is to be filtered and compressed by orders of magnitude. One challenge is to define these filters in such a way that they do not discard useful information. Research is needed in the area of data reduction that can intelligently process raw data to a size that its users can handle without losing essential information. Furthermore, on-line analysis techniques that can process streaming data on the fly are required for Big Data.

The second big challenge is to automatically generate the right metadata to describe what data is recorded and how it is recorded and measured. For example, in scientific experiments, considerable detail regarding specific experimental conditions and procedures may be required to interpret the results correctly, and it is important to record such metadata along with observational data. Metadata acquisition systems can minimize the human burden in recording metadata. Another important issue here is data provenance. Recording information about the data at its birth is not useful unless this information can be interpreted and carried along through the data analysis pipeline. Thus, research is needed in both directions of generating suitable metadata into data systems and carrying the provenance of data and its metadata through data analysis pipelines.

### B. Information Extraction and Cleaning

The information collected will not be in a format suitable for analysis. For e.g., the collection of electronic health records in a hospital comprises of transcribed dictations from several physicians, structured data from sensors and measurements, and image data such as x-rays. This data cannot be effectively analyzed. An information extraction process is needed to extract required information from the underlying sources and to express it in a structured form suitable for analysis. Doing this correctly and completely is a continuing technical challenge. Data extraction is often highly application dependent. In addition, due to the ubiquity of surveillance cameras and popularity of GPS-enabled mobile phones, cameras, and other portable devices, rich and high fidelity location and trajectory (i.e., movement in space) data should also be extracted.



Big Data are not trust worthy. Existing work on data cleaning assumes well-recognized constraints on valid data or well-understood error models; for many emerging Big Data domains these do not exist.

### C. Data Integration, Aggregation and Representation

Given the heterogeneity of the flood of data, it is not enough to record it and store into a repository. Data analysis is considerably more challenging than simply locating, identifying, understanding, and citing data. For effective large-scale analysis all of this has to happen in a completely automated manner. This requires different data structures and semantics to express data in forms that are computer understandable, and then “robotically” resolvable. In spite of some basic work in data integration, additional work is required to achieve automated error-free difference resolution.

Selecting a suitable database design to store information is another challenge. Effective database designs should be created, tools should be devised to provide assistance in the design process, or techniques should be developed to use the databases effectively even in the absence of intelligent database design.

### D. Query Processing, Data Modeling and Analysis

Methods for querying and mining Big Data are fundamentally different from traditional methods. Big Data is noisy, dynamic, heterogeneous, inter-related and not trustworthy. Even noisy Big Data could be more valuable because general statistics obtained from frequent patterns and correlation analysis usually overpower individual fluctuations and disclose more reliable hidden patterns and knowledge. Further, interconnected Big Data forms large heterogeneous information networks, with which information redundancy can be explored to compensate for missing data, to crosscheck conflicting cases, to validate trustworthy relationships, to disclose inherent clusters, and to uncover hidden relationships and models.

Data mining can be used to improve the quality and trustworthiness of the data, understand its semantics, and provide intelligent querying functions. Big Data is enabling the next generation of interactive data analysis with real-time answers. In the future, queries towards Big Data will be automatically generated for content creation on websites, to populate hot-lists or recommendations, and to provide an ad hoc analysis of the value of a data set to decide whether to store or to discard it. Scaling complex query processing techniques to terabytes while enabling interactive response times is a major open research problem today.

A problem with current Big Data analysis is the lack of coordination between database systems, which host the data and provide SQL querying, and analytics packages that perform various forms of non-SQL processing (such as data mining and statistical analyses). Today’s analysts are impeded by a tedious process of exporting data from the database, performing a non-SQL process and bringing the data back. This is an obstacle for availing the interactive elegance of the first generation of SQL-driven OLAP systems, in the data mining type of analysis, with increasing demand. A tight

coupling between declarative query languages and the functions of such packages will benefit both expressiveness and performance of the analysis.

### E. Interpretation

Big Data Analytics is of limited value if users cannot understand the results. Results interpretation involves examining all the assumptions made and retracing the analysis. This is particularly a challenge with Big Data due to its size and complexity. Analytical pipelines can often involve multiple steps, again with assumptions built in along with crucial assumptions in data recording.

Providing results is not enough. One must also provide supplementary information that explains how each result was derived, based upon what inputs. Such supplementary information is called the provenance of the (result) data. By studying how best to capture, store, and query provenance, in conjunction with techniques to capture adequate metadata, an infrastructure can be created to provide users with the ability both to interpret analytical results obtained and to repeat the analysis with different assumptions, parameters, or data sets.

Systems with a rich palette of visualizations are important in conveying the results of the queries to the users in a way that is best understood in the particular domain. Furthermore, users require not only the results, but also should understand why they are seeing those results. However, raw provenance, particularly regarding the phases in the analytics pipeline, is likely to be too technical for many users to grasp completely. One alternative is to enable the users to “play” with the steps in the analysis – make small changes to the pipeline, or modify values for some parameters. The users can then view the results of these incremental changes. By these means, users can develop an intuitive feeling for the analysis and also verify that it performs as expected in corner cases. Accomplishing this requires the system to provide convenient facilities for the user to specify analyses.

## V. TECHNICAL CHALLENGES

Technical Challenges that underlie all phases of Big Data Analytics Pipeline follow:

### ➤ Scale

Managing large and rapidly increasing volumes of Big Data raise challenges. A drift in processor technology to intra-node parallelism and packing multiple sockets; move towards cloud computing and transformative change of traditional storage subsystems from Hard Disk Drives to solid state drives and other technologies such as, Phase Change Memory necessitates advances in the storage and processing of Big Data growing at an unprecedented scale day-by-day.

### ➤ Timeliness

The flip side of size is speed. Velocity in the context of Big Data imposes a timeliness challenge as the result of the analysis is required immediately in many situations. New index structures are required to speed up the search in

situations such as, finding fraudulent credit card transactions. Designing such structures become particularly challenging when the data volume is growing rapidly and the queries have tight response time limits.

### ➤ *Heterogeneity and Incompleteness*

To enable transformative opportunities, information from different multiple data sources should be integrated. Handling and analyzing such data poses several challenges as the data can be of different types: structured, semi-structured, unstructured or mixed. Different data sources may use their different formats. To complicate matters further, data can arrive and require processing at different speeds: batch, near-time, real-time or streams. Standard formats and interfaces are required to deal the issues of heterogeneity.

Some incompleteness and some errors are likely remain in data, even after data cleaning and error correction, This incompleteness and these errors must be managed during data analysis. Doing this correctly is a challenge.

### ➤ *Privacy*

Privacy is an issue whose importance is growing as the value of Big Data becomes more apparent. In addition, real data is not static but gets larger and changes over time,; none of the prevailing techniques result in useful content being released in this scenario. Personal data such as, health records of financial records offer most significant human benefits and are most sensitive at the same time. Trade-offs between privacy and utility is to be essentially taken care of while designing an effective Big Data strategy.

Another closely related concern is data security. With serious breaches on the rise, addressing data security through technological and policy tools became essential. Yet other very important direction is to rethink security for information sharing in Big Data use cases and social media. Managing privacy is effectively both a technical and sociological problem that must be addressed jointly from both perspectives to realize the potential of Big Data.

### ➤ *Human Collaboration*

In spite of the tremendous advances made in computational analysis, there remain many patterns that humans can easily detect but computer algorithms have a hard time finding. A Big Data analysis system must support input from multiple human experts may be separated spatially and temporally and shared exploration of results. The data system has to accept this distributed expert input, and support their collaboration. The issues of uncertainty and error become even more pronounced with Big Data. The extra challenge here is the inherent uncertainty of the data collection devices. Technologies are needed to facilitate detection and correction of false information provided and a framework is also necessary for analysis of data with conflicting statements and to assess their correctness.

The very fact that Big Data analysis typically involves multiple phases highlights a challenge that arises routinely in practice: production systems must run complex analytic

pipelines, or workflows, at routine intervals, e.g., hourly or daily. New data must be incrementally accounted for, taking into account the results of prior analysis and pre-existing data. And of course, provenance must be preserved, and must include the phases in the analytic pipeline. Current systems offer little to no support for such Big Data pipelines, and this itself is a challenging objective.

## EXECUTIVE SUMMARY

Data have become a torrent flowing into every area of global economy. The term 'Big Data' is coined to describe this Information Tsunami which impacts every aspect of business operations virtually in every sector. 'Big Data Analytics' describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and/or analysis, while ensuring their veracity by an automatic quality control in order to obtain a value. The effective use of Big Data has the potential for transforming economies to deliver a new wave of productivity growth and consumer surplus. Big Data will help to pinpoint right medical treatment, financial product or service.

Despite the hype, Big Data Analytics is still a challenging, complex and time demanding endeavor, requiring expensive software, computational infrastructure and effort. There is a shortage of talent necessary for organizations to take advantage of Big Data. Several issues such as, data policies, technology & techniques, infrastructure, provenance, privacy, heterogeneity, scale, timeliness, visualization will have to be addressed to tap into the full value of Big Data. Furthermore, these challenges require transformative solutions.

The potential opportunities of Big Data should be assessed against the strategic threats and any gap between the current IT capabilities and data strategies to capture big data opportunities relevant to consumer or enterprise should be closed. Creative and proactive strategies should be developed for combining right pools of data to create value and to gain access to those pools, with security and privacy issues as well.

The research should focus on helping consumers to understand the benefits of the Big Data along with the risks. In parallel, the research should proceed in a direction to design and develop tools, techniques and technologies that are data savvy.

This paper presents a brief review on Big Data and Analytics with a focus on Applications, Phases of analytics pipeline and Technical Challenges encountered in each of these phases. To conclude, 'Big Data and Analytics' is a research frontier for Innovation, Competition and Productivity.

## REFERENCES

- [1] A community white paper developed by leading researchers across the United States, "Challenges and Opportunities with Big Data"
- [2] Abdelkader Baaziz and Luc Quoniam, "How to use Big Data technologies to optimize operations in Upstream Petroleum Industry"
- [3] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs Charles Roxburgh, Angela Hung Byers, "Big Data; The next frontier for innovation, competition and productivity", McKinsey & Company, May 2011
- [4] L. Venkata Subramaniam, "Big Data and Veracity Challenges", Text Mining Workshop, ISI Kolkata, Jan 8, 2014
- [5] Marcos D. Assuncao, Rodrigo N. Calheiros, Silvia Bianchi, Marco A. S. Netto, Rajkumar Buyya, "Big Data Computing and Clouds: Challenges, Solutions, and Future Directions", arXiv:1312.4722v1 [cs.DC] 17 Dec, 2013
- [6] Omer Tene\* & Jules Polonetsky, "Privacy in the Age of Big Data: A Time for Big Decisions", Symposium Issue, 64 Stan. L. Rev. Online 63, 2 Feb, 2012
- [7] Research Trends Special Issue on Big Data, 30 Sep, 2012

**Authors**

Ms. S. Suneetha received her Bachelor's Degree in Science and in Education, Master's Degree in Computer Applications (MCA) from SVU, Tirupati and M.Phil. in Computer Science from SPMVV, Tirupati. Currently, she is pursuing her Ph.D. in SPMVV, Tirupati. She is a life time member of ISTE.

Her areas of interest are Data Mining & Software Engineering. She has 15 papers in National/International Conferences/ Journal to her credit. She also attended several workshops in different fields. She served Narayana Engineering College, Nellore, Andhra Pradesh as Sr. Asst. Professor, heading the depts. of IT & MCA.



Dr. M. Usha Rani is Professor in the Department of Computer Science, Sri Padmavati Mahila Viswavidyalayam (SPMVV Women's' University), Tirupati. She did her Ph.D. in Computer Science in the area of Artificial Intelligence & Expert Systems.

She is in teaching since 1992. She presented many papers at National and International Conferences and published articles in National & International journals. She has also written 4 books like Data Mining - Applications: Opportunities and Challenges, Superficial Overview of Data Mining Tools, Data Warehousing & Data Mining and Intelligent Systems & Communications. She is guiding M.Phil. & Ph.D. Scholars in areas like Artificial Intelligence, DataWarehousing & Data Mining, Computer Networks and Network Security, Cloud Computing etc.