

A COMPREHENSIVE STUDY ON POTENTIAL RESEARCH OPPORTUNITIES OF BIG DATA ANALYTICS TO LEVERAGE THE TRANSFORMATION IN VARIOUS KEY DOMAINS

Dr. V.V.R. Maheswara Rao¹, Dr. V. Valli Kumari² and N. Silpa³

¹Professor, Department of Computer Science & Engineering,
Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India,

²Professor, Department of Computer Science & Systems Engineering,
College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India,

³Asst. Professor, Department of Computer Science & Engineering,
Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India,

ABSTRACT

Companies, organizations and policy makers shake out with flood flowing volume of transactional data, accumulating trillions of bytes of information about their customers, suppliers and operations. The advanced networked sensors are being implanted in devices such as mobile phones, smart energy meters, automobiles and industrial machines that sense, generate and transfer data to multiple storage devices. In fact, as they go about their business and interact with individuals, they are producing an incredible amount of fatigue digital data. Social media sites, smart phones, and other customer devices have allowed billions of individuals around the world to contribute to the amount of data available. In addition, the extremely increasing size of multimedia data has also take part a key role in the rapid growth of data. The technology of high-definition video creates more than 2,000 times as many bytes as necessary to store as normal text data. Moreover, in a digitized world, consumers are leaving enormous amount of data about their day-to-day communicating, browsing, buying, sharing, searching and so on. As a result, it evolved as a big data and in turn has motivated the advances in big data analytics paradigms, endorsed as a basic motivation factor for the present researchers.

The authors in the present paper conduct a comprehensive study to explore the impact of big data analytics in key domains namely, Health Care (HC), Retail Industry (RI), Public Governance (PG), Pubic Security & Safety (PSS) and Personal Location Tracking (PLT). Initially, the study looks at the insights of data sources along with their characteristics in each domain. Later, it presents the highly productive and competitive big data applications with innovative big data technologies. Subsequently, the study showcases the impact of big data on each domain to capture value addition in its services. Finally, the study put forwards many more research opportunities as all these domains differ in their complexity and development in the usage of big data analytics.

KEYWORDS

Big Data, Big Data Analytics, Big Data Technologies, Big Data Applications

1. INTRODUCTION

Recent technological advancements collect more data than ever before and led to a torrent of data from individual domains over the past two decades. The size of the data, evolved into its present stage as mega bytes to giga bytes, giga bytes to terra bytes, terra bytes to peta bytes, peta bytes to exa bytes as shown in figure 1. The idiom big data was created to confine the meaning from this mountain of data. In addition to its huge volume, big data has unique characteristics while compared with conventional data. Extracting useful real time information on such enormous data has evolved as solid base for the present researchers to enter into the era of big data analytics.

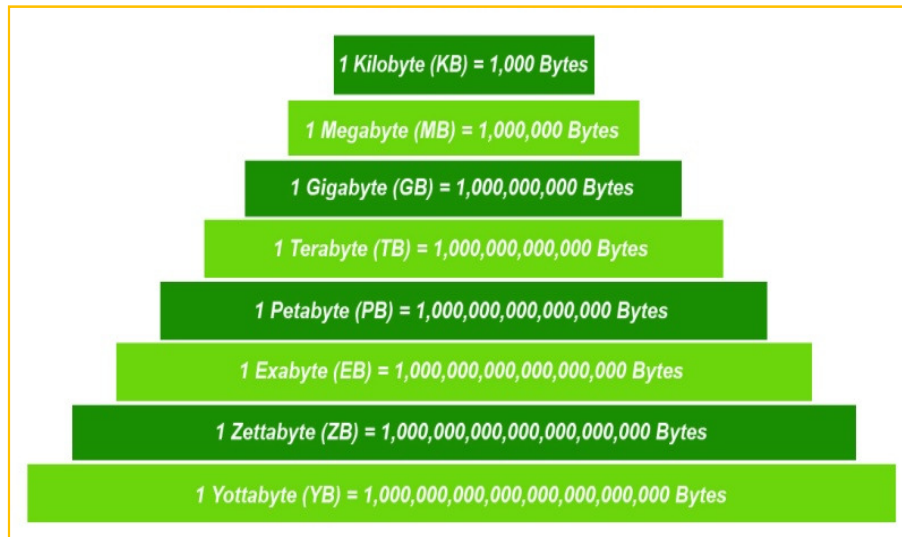


Figure 1. Mountain of Data

The rapid advancements in digitization, the amount of data being generated from various resources have reached astronomical proportions. This data is structured, semi structured, unstructured, complex, heterogeneous, high dimensional and incremental in nature as shown in figure 2. Big data not only means a fundamental shift in the way data is stored and managed, it also entitles deploying powerful real time analytics and visualization tools, collaboration platforms and the ability to automatically create links with the existing applications such as business support systems and customer relationship management.

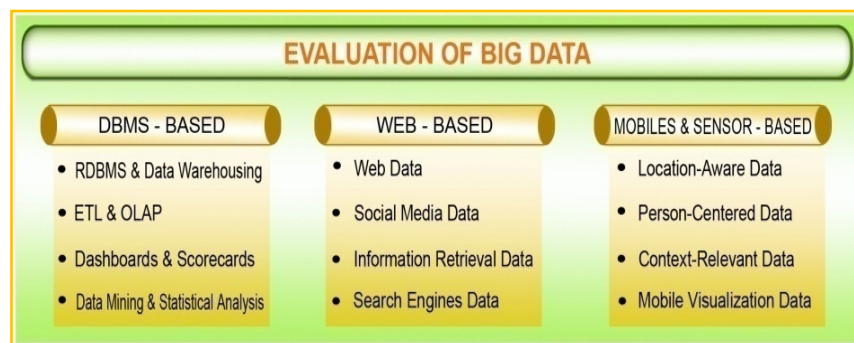


Figure 2. Evaluation of Big Data

In technology perspective, big data is the possibility of better storage -Volume, the ability to process the information and make it available in real time - Velocity and the ability to deal with various kinds of data sources, including structured, semi-structured and unstructured ones - Variety. Inclusion of Veracity as the fourth big data attribute emphasizes the importance of addressing and managing for the uncertainty inherent with in some times of data. The technology exists, so the essential issue is how carriers can make sense of the massive volumes of data and deliver as fifth attribute Value to business.

Fundamentally, big data means not only 5V's of data but also describes a new generations of technologies and architectures, design to economically extract value from very large volumes of a wide variety of data, by enabling velocity capture, discovery and analysis. This definition converge the five dimensions: Volume, Velocity, Variety, Veracity and Value as shown in figure 3 help both to define and describe big data.

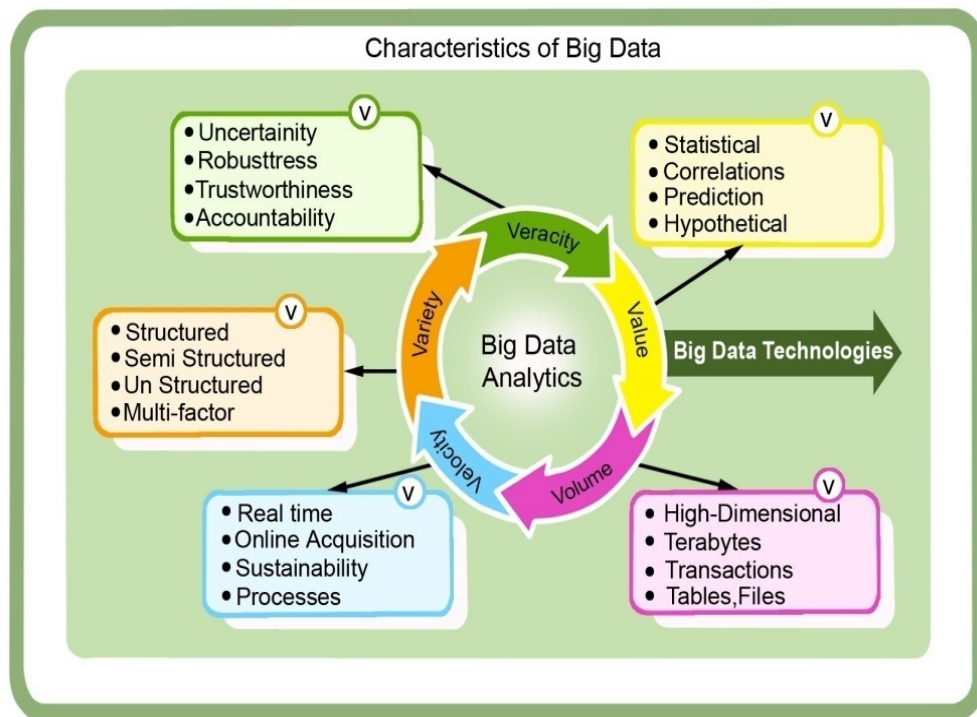


Figure 3. Characteristics of Big Data

Volume: Volume indicates the size of the data, the volume of big data evolved into its present stage as mega bytes to giga bytes, giga bytes to terra bytes, terra bytes to peta bytes, peta bytes to exa bytes. The big data volumes are relative and vary by factors, such as time and the type of data. This data is too big to be handled by the current state of techniques and systems. In future, this will continue to expand exponentially at an unprecedented rate, is a prime motivation to create revolutionary data management mechanisms.

Velocity: Velocity refers to the rate at which data are generated and the speed at which it should be analyzed and acted upon. Velocity in big data is a concept which processed and analyzed with the speed of the data coming from various sources. The proliferation of digital devices such as smart phones and sensors has led to an unprecedented rate of data which is continually being generated at a pace that is impossible for traditional systems to capture, store and analyze. This,

coupled with the need and drive to be more agile and deliver insight quicker, is a basic motivation for the present data engineering researchers to build the necessary infrastructure and skill base to react quickly.

Variety: Variety shows different types of data and data sources. Variety in big data is a measure of heterogeneity of data representation such as structured, semi structured and unstructured. With the explosion of sensors, smart devices and social collaboration technologies, data is being generated in countless forms, including: text, web data, tweets, sensor data, audio, video, click streams, log files and more. However, the emergence of new data resources enables the researchers towards new management technologies and analytics, which enable to leverage data in an innovative aspect.

Veracity: Veracity denotes data uncertainty. Veracity in big data is the level of reliability associated with certain types of data. Some data is inherently uncertain, for example: sentiment, truthfulness, weather conditions, economic factors. The need to acknowledge and plan for this dimension of uncertainty of big data is still a major quality concern in processing of data. This triggered the present researchers towards the robust optimization techniques to manage uncertainty.

Value: A new generation of big data technologies and architectures designed to economically extract value from very large volumes of data.

In the present paper the authors study on the impact of big data analytics in various domains such as Health Care (HC), Retail Industry (RI), Public Governance (PG), Public Security & Safety (PSS) and Personal Location Tracking (PLT) to explore research challenges and opportunities. This paper is organized as follows. Section 2 describes related work. Later, section 3, presents proposed research study in five key domains. Finally, in section 4 conclusions and future work are mentioned.

2. RELATED WORK

The comprehensive study in the present paper contains literature survey from 2010 to current year. In specific, the authors in the present study concentrate five applicable key domains: health care, retail industry, public governance, public safety & security and personal location tracking and explored the complete research challenges & opportunities extensively and comprehensively. Towards this the research review and the contributions made by the authors [31, 32, 34] noticed and emphasises that the big data analytics includes interdependent layers namely big data resources, big data technologies, big data applications, and impact of big data. Furthermore, the recent literature study [1, 5, 6] has recognized that implementation of interdependent layers comprehensively is a promising and practical research area.

The authors, [21, 36] explained the evaluation changes of digital world and their causes in 21st century. Later they presented the big data pipeline and list out the challenges in big data. They finally described that big data analytics as an emerging type of knowledge work, with plenty of opportunities in different domains. In addition, some more authors [28, 33] have attempted to offer a broader definition of big data and its characteristics. They primarily focused on the analytic methods to leverage massive volumes of heterogeneous data in unstructured text, audio and video formats. However, they worthily emphasized that real world adoption of big data analytics were not economically feasible for large scale applications. Also, in their conclusion, they felt that novel big data analytics not yet taken place and it becomes a prolific field of research.

Subsequently [1, 5, 30] developed a report based on analysis of survey data and discussions with academics, subject experts, business executives. They mainly concentrated on the concept, characteristics, sources, analytics capabilities, adoption stages and primary obstacles of big data. In their conclusion, they clearly suggested that future research demands the effective use of information and analytics to understand comprehensive needs of the organizations.

In the health care domain, [8, 12, 15, 25, 35] wrote a report on the promise and peril of big data. They provided the basic concepts of big data and explained business and social implications of big data. They have expressed that medical researchers focus on to identify useful correlations between medical treatments and health outcomes, that helps to improve health and medical care can be made more efficient and effective. The authors [3, 8] undoubtedly expressed big data effects heavily on overall patient care, business of healthcare delivery, healthcare regulation, transparency & justification and operational efficiency & profitability in services significantly. Some more authors [2, 26] in their study reported that the major challenges with a focus on three upcoming and promising areas of medical research: image, signal, and genomics based analytics. In the retail industry domain, [9, 10, 15] addressed key technologies for extracting real-time business value from the big data. They summarized that only one-third of organizations do big data analytics, the present research given the newness of the combination of advanced analytics and big data. A industry-specific report [13, 18, 19, 27, 29] is generated and they examined the retail industry respondents view towards big data and to what extent it is currently using to benefit of the businesses. They also felt that big data analytics makes a cultural shift in the way retailers connect with consumers in a meaningful way in transforming their processes, organizations and, soon, the entire industry.

In the public governance sector, the authors [3, 4, 7] recognized that big data analytics derive the value from massive amounts of digital data. They also reported that governments of leading ICT countries have initiated big-data application projects to enhance operational efficiency, transparency, citizens' well-being and engagement in public affairs, economic growth, and national security. In addition they suggest that analyzing big-data application projects by governments offers guidance for follower countries for their own future big-data initiatives. The survey [14, 16, 23, 24] clearly shown that most public sector agencies are struggling to deliver better service to their citizens while contending with severe financial constraints. But a host of technology innovations can ensure safety, further improve quality of life and increase confidence in government.

In the public security and safety, the authors [11, 30] suggests that awareness of data collection and use by government is quite high, but that the level of understanding is much lower. They also suggest that perceptions are likely to change as understanding about potential implications increases. In addition they notified there is limited research specifically on the use of big data for law enforcement & intelligence and personal privacy concerns. Some more authors [5,17] reported that advanced analytics extracts insights from public safety and security big data, supports the analyst beyond the capability to just ask specific questions or run specific queries. The analytical techniques will model the data and push information of interest back to the user. Multiple analytical techniques are applied to the large data volumes to uncover the key nuggets of intelligence.

In personal location tracking domain, the authors [20, 21] identified that cost minimization has become an emergent issue for the upcoming big data era. Different from conventional cloud services, one of the main features of big data services is the tight coupling between data and computation as computation tasks can be conducted only when the corresponding data are available. In addition they reported that factors, like task assignment, data placement and data

movement, deeply influence the operational expenditure of data centers. Furthermore, they extensively simulated to get the high efficiency of their proposal. In addition, they revised the opinion of big data and presented the new view of big data. They concluded that the researchers have to pay attention not only on the characteristics of big data should also work on analytics of big data.

Many of the earlier authors as observed in the literature have explained the importance, criticality and efficiency of comprehensive approach in the process of big data analytics, which has been considered as the formal basis for the present study.

3. PROPOSED RESEARCH STUDY

To explore the complete research challenges & opportunities and recognize the impact of big data analytics at each stage, the authors identify five broadly applicable key domains namely, Health Care (HC), Retail Industry (RI), Public Governance (PG), Public Security & Safety (PSS) and Personal Location Tracking (PLT) as shown in figure 4.

Towards this, in each domain, initially, the study focus on growing torrent big data resources and looks at the insights of characteristics of data to explore the major challenges. Later, it concentrates on big data technologies to extract knowledge from such resources. Subsequently, pay an attention on to recommend intelligence systems that are highly productive and competitive of big data applications. Finally, the study, enlighten the impact of big data to capture its value addition in its services.

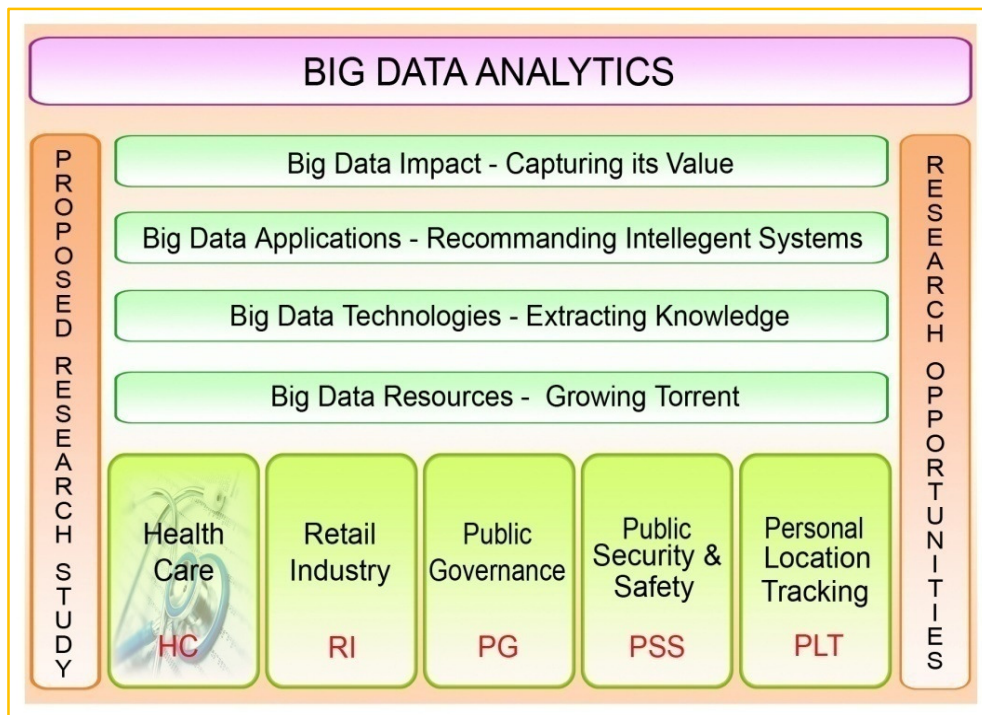


Figure 4. Block Diagram of Proposed Research Study

3.1 Health Care (HC)

Health care is a huge and significant segment of the world economy that faces great productivity problems. Health care domain has many stakeholders, including patients, doctors, pharmaceutical products industries, medical product industries and so on. These stake holders have diverse interests and needs while still being strongly intertwined. The characteristic of health care data are structured, semi-structured, un-structured and even a significant portion of clinical data is not yet digitized. There is a considerable opportunities to create value if this data is digitized, combined and used successfully.

3.1.1 Big Data Resources in Health Care

The health care big data creates from several patient care points of contact, sophisticated medical apparatus, web-based health portals and social media networks. The major resources of health care data are genomics-driven data in form of genotyping, gene expression, sequencing data and payer-provider data in form of electronic patient health records, insurance policy records, medical prescription, patient feedback forms and responses.

Extracting knowledge from such health care big data poses significant research and practical series of challenges. The possibilities of health care big data continue to evolve rapidly and have reached astronomical proportions as it collected from various resources of hundreds of thousands of patients with many reports and inherent uncertain factors over a lengthy period of time. The health care data is sensitive in nature as it is person specific content and disparate with highly linked way. Developing a privacy and reliable health infrastructure and accomplishing an ethical health data based research is another open challenge.

3.1.2 Big Data Technologies in Health Care

An ample range of technologies has been developed and tailored to aggregate, manipulate, analyze and visualize health care big data effectively. All the technologies embedded the principals of flexible and interdisciplinary areas including statistics, applied mathematics, computer science techniques and economics to create value addition. Any such technology ensures to extract noteworthy clinical knowledge, profound understanding of patient disease patterns and associations in medical diagnoses symptom based diseases from health care big data efficiently.

There is a need for ongoing innovation in technologies and techniques such as association rule mining and clustering, health social media monitoring and analysis, health text analytics, health ontologies, patient network analysis, genomics and sequence analysis, privacy preserving data mining and adverse drug side-effect analysis to extract knowledge health care data are future promising areas of research.

3.1.3 Big Data Applications in Health Care

The big data applications in the health care data are the interfaces provided by the big data technologies that can substantially recommend intelligent systems to improve decision making, minimize risk, unearth valuable insights and so on by performing sophisticated analysis in a timely manner.

Towards this, health care decision support system, patient community analysis, medical procedure analysis, transparency systems, remote patient monitoring systems, advanced patient

profile analytics, human genomic analysis are the emerging next leading edge applications in health care domain and create another trend research opportunities.

3.1.4 Big Data Impact in Health Care

The impact of health care big data critically enables a range of levers which can reduce the cost, improve the quality in health care, efficient quality of services in long term care and patient empowerment. In addition to that it identifies performance opportunities for medical professionals, processes and organizations. Moreover, the impact monitors chronically ill patients to improve future drug and treatment options. It also identifies patient segments that would benefit from proactive care or lifestyle changes and would benefit from a preventive care programs. It is worth emphasize that big data analytics radically regulates the policies and procedures. Furthermore, health care domain achieves a significant value available at insights of big data.

3.1.5 Summary on Health Care

The health care domain faces an urgent imperative to attain an essential improvement in its cost-effectiveness. This study shows that a range of parameters using big data enables critically in reducing cost and maintaining health care quality. However, this domain will not achieve the significant value available from big data without radical changes in regulations and systems. The power of big data empowered the health care domain with the following outcomes as shown in figure 5.



Figure 5. Outcomes of Big Data in Health Care Domain

3.2 Retail Industry

Retail Industry is also a rich domain, the interactions among retailers and consumers plays a significant value in the global economy as there is a close relationship with market economics as a whole. The retail industry has its own sub divisions like, general merchandise, building materials & garden equipment, non-store retailers, food & beverage, clothing & accessories, sporting goods, books, electronics & appliances, and furniture & home furnishings. This is an area in which digital data is playing an escalating role as consumers search, research, compare, buy and obtain support online, and the products sold by retailers generate their own digital data progressively more.

As a result, it faces many challenges to enhance the productivity, to face competitiveness of companies and to create economic surplus for consumers. Thus, in the retail industry, the use of

information technology and digital data has been instrumental in boosting the profitability of individual players and the productivity of the entire sector for decades. Hence, in this domain the continued adoption and development of big data to overcome barriers related to technology, talent and organizational culture.

3.2.1 Big Data Resources in Retail Industry

A variety of new technologies disseminates the data at an unprecedented rate in retail industry. Data sources such as historical customer information, customer transaction data, a consumer's search & user logs, mobile interaction, CRM data, credit card usage, in store video data and a vast amount of unstructured social media informal customer opinion data is available. Volume of this data is growing inevitably as retailers not only record every customer transaction but also keep track of emerging data sources such as barcode, RFID chips, catalogs that track products, as well online customer behaviour.

To overcome multiple barriers and realizing the full potential of retail industry from such diverse and complex resources and the velocity with which it is growing is turn into a great research challenge. The research has to pay an attention mining peta bytes of customer data, customer preferences and buying behaviour that can enable retailer to improve the effectiveness of their marketing and average purchase size. In addition to that it also offers long term competitive advantages and strategies for growing retailers' revenue.

3.2.2 Big Data Technologies in Retail Industry

The retail industry powered with a great deal of legacy IT systems in order to capture value from data. These conventional techniques usually generate information in incompatible standards & formats and pose great challenges in integrating, accepting and analysing. But, the present demand of the retail industry focuses on scalable and integrated systems with advanced key technologies in data analysis. There is also a need for ongoing innovation in technologies and techniques that will help retailers to integrate, analyze and visualize the growing torrent of consumers data.

Now a days, the amount of unstructured data is larger than structured data in retail industry. However, the top level management needs both the data to make right decisions in getting avenues in the commutative business. To attain this, future of information management in this industry requires a new brand of technologies and tools have the capability of process of high volume & velocity of unstructured data, enable traditional online processing and handle high veracity of real data quickly. In this situation the accuracy of retrieving the knowledge through big data technologies such as database segmentation, association rule mining, clustering, anomaly detection, social network analysis, text and web analytics is a potential open challenge.

3.2.3 Big Data Applications in Retail Industry

The big data applications in the retail industry transform the business in a holistic way through reduced costs and improved operating margins. These applications derive greater insights of data with real time capability. They also provide transformative research opportunities in each key parameter of identified categories like marketing, merchandising, operations and supply chain. Moreover, these applications enable the retail industry with high success rate and create a fundamental change in all such categories. In addition, they allow the retailers to make instantaneous decisions, thus improves customer satisfaction.

The application in the retail industry open a wide range of research opportunities in micro segmentation, price optimization, inventory management, in-store analytics, location-based

marketing, generating customer loyalty, predictive demand, performance transparency, labour input optimization and web based markets. Indeed the research also suggests towards the tremendous wave of innovation, productivity and growth as well as new modes of competition to exploits potential value in retail industry.

3.2.4 Big Data Impact in Retail Industry

Big data enables better and granular customer segmentation by collating customer information from various sources. In addition, it optimises inventory, minimises inefficiencies and avoid stock outs. It also ensures better cross sell and up sell capabilities. The big data analytics provides 360 degree customer view by leveraging unused data assets. Thus, the transformative nature of big data drives the retail industry with true business value.

3.2.5 Summary on Retail Industry

Retail industry faces significant pressures from suppliers upstream and consumers downstream to ensure their survival and relevance. While information technology and data have already delivered waves of impact, this study finds that there is significant headroom arising out of innovation that retailers can tap. Retailers that develop and exercise their big data muscles will boost their outcomes as shown in figure 6.



Figure 6. Outcomes of Big Data in Retail Industry

3.3 Public Governance

Public governance in countries across the globe is under heavy pressure to enhance productivity of their services. It has many multilevel diversified administrative functions and has high level responsibility of services with significant budgetary constraints. The effective governance of these administrative functions turn into a great challenge as its functions deeply dependant horizontally and vertically to offers services. This situation causes the public governance systems with typical difficulties in quality, flexibility and effectiveness of public services. These circumstances exceed the capability of traditional information technology support to integrate and evaluate information assets. Thus, there is a potential research opportunities in the adaption of power full big data concepts to drive the public governance innovatively.

3.3.1 Big Data Resources in Public Governance

With the improved usage of affordable & faster computer and communication systems in public governance, the rate of digital data has been increased exponentially. Moreover, the low cost

digital storage has created another opportunity for increased data. In addition, the wide usage of disruptive technologies such as mobile devices, cloud services and social media technologies in high value public governance systems creates massive amount of data. Also, the small handheld devices and applications have significantly contributed in the generation of digital data. The e-governance initiatives in public governance is another channel to generate high volumes of the digital.

The vast majority digital data available in public governance legacy systems, collected from multiple channels and sources is in structured, semi structured, unstructured and cohesive formats. Extracting useful information from such data by the legacy systems is unable to generate standardized solutions. This situation opens avenues for big data researchers with many more research opportunities.

3.3.2 Big Data Technologies in Public Governance

A wide range of technologies has been incorporated by public governance towards data capturing, storing and analysing the efficiency of government services. Towards this, the public governance made an attempt to incorporate transparency and accountability of services by establishing traditional tools and infrastructure. But, the larger and more varied digital data coming at high velocity in this sector made traditional tools and technologies inefficient and non scalable in analysing the data. It is evidence that productivity growth of public governance sector has fallen behind that of private sector in many parameters.

To increase the transparency, scalability and productivity in managing such data in the public governance necessitate the emerging of big data technologies. These emerging technologies majorly focus on integration of multi formatted data, text analytics, analysis of semantic services, monitoring of social media informal data and analysis on sentiment data. In addition, rule based algorithms and well tuned algorithmic techniques such as neural networks also useful in public governance. The outcome of these new technologies in this domain aimed to improve transparency and scalability, boosting productivity, achieving higher levels of efficiency, reduce the cost of administrative activities promotes many more promising research paths.

3.3.3 Big Data Applications in Public Governance

The big data applications are able to conduct micro analysis on the public governance data and also discover specific insights in the data. The applications help to raise the public governance administrative activities by incorporating incremental and intelligent systems effectively. In addition, they accelerate annual productivity growth & efficiency gain of public governance systems.

The applications mainly enable public governance to create enhanced transparency, ability to explore the performance, effective customized segmentation, automotive decision making and innovative management practices etc. Additionally, they allow the public governance to conduct feedback on services and policies. However, the governance system expect big data applications to enhance the ability in order to address major challenges involving transparency, economy, health care, employment, natural disaster and public security. Thus, these challenges create many more research opportunity in micro analysis, cost optimization, generating transparency, equal access of public services, public participation and provide better services to human kind.

3.3.4 Big Data Impact in Public Governance

Big data analytics influence the public governance sphere that includes public service provisions, public service organizations, public participation, e-governance systems and interdepartmental collaborations that transform the role of government and its operational efficiency. It also emphasizes on changing the fundamental nature of relationship between governance systems and citizens. The big data analytics able to poster online direct democracy in the administrative, political and legislative front. The big data infrastructure and intelligent citizen interfaces raise the public governance to global competitive as it attracts public talent into the governance. The transparency nature of big data analytics improves private and public partnership. Thus, big data creates value for public service delivery to all stakeholders.

3.2.5 Summary on Public Governance

The public governance domain in advanced economies is facing unprecedented pressures to improve productivity to make a significant contribution around the world. Using big data intelligently has the potential to generate significant outcomes are showcased in figure 7.



Figure 7. Outcomes of Big Data in Public Governance Domain

3.4 Public Security and Safety

Public security and safety system is of high maturity domain, with a large amount of data accumulated from service & social resources and struggle with lot of challenges to retrieve intelligent and relevant information. This domain is closely associated with officers, investigators, intelligence officials working at law enforcement agencies, security agencies, national level databases, national crime agencies, secret intelligent agencies and general public. The entire stakeholders involved in such agencies have high expanded requirements for timeliness, completeness & accuracy of information in practice. Ability to understand the root causes of nature of crimes across time, location and many other factors catch the attention of the big data researchers and open wide range of research opportunities.

3.4.1 Big Data Resources in Public Security and Safety

In public security domain, lot of valuable data is in criminal records, crime maps, criminal networks, news and web content, terrorism incidents databases, viruses and cyber attacks that are available in unstructured formats such as documents, transcripts and witness statements. Integrating, processing and extracting knowledge from such data struggled with many typical challenges. In addition the hidden high potential and great value data requires human intervention to make the sense.

Big data analytics having sophisticated linguistic rules and statistical methods that automatically determine, categorize and manage semantic terms are very much required in this domain. Specifically, databases in this domain hold criminal records, DNA profiles, finger prints and data from local agencies, create technical challenges to maximize the usage of big data. As all stakeholders concentrating on witness description that contains most valuable information, holistically big data opens more research paths in this front.

3.4.2 Big Data Technologies in Public Security and Safety

The number of technologies has been adopted in this domain to retrieve the information to serve the stakeholders. All these technologies inspired by the principals of patterns, correlations, predictive policies and associations. However, all these traditional technologies extract the knowledge with limited evidence and leads to failure predictions. In addition, the traditional computing architecture is not easy to expand infrastructure capacity and faces difficulties in deep mining with high volumes of varied data with high velocity.

There is a definite need of innovation in technologies such as criminal association rule mining, clustering, criminal network analysis, special temporal analysis, text analysis, sentiment analysis, cyber attacks analysis, and visualization to process data effectively with high requirements of visual display. This situation creates a future promising research in public security and safety domain with big data technologies as whole and individually in wide area.

3.4.3 Big Data Applications in Public Security and Safety

The big data applications in public security and safety make over the analysis and able to derive the sense from unstructured data with less human intervention in a comprehensive way to improve consistent intelligence. The applications majorly look for new insights from the routinely collect large amount and type of data and highlight areas for investigation in order to prevent the occurrence of crime so as to provide the safety and security to the people. In addition, they address security challenges timely and reliable intelligence. Moreover, these applications facilitate all the stakeholder of public security and safety domain to extract intelligence from unstructured data in conjunction with structured data. This intelligence helps the domain to ensure real time practical needs of all intelligence operations.

Big data applications in public safety and security create an extensive opportunities in crime analysis, computational criminology, terror informatics, open source intelligence, cyber security, crime prevention, crime detection and national security. In reality the research also put forward a comprehensive model to explore real intelligence to regulate public issues like privacy, civil liberty and social benefits.

3.4.4 Big Data Impact in Public Security and Safety

The usage of big data in public security and safety potentially exploit the areas to occur crime, regularity in the occurrence and also informing strategic planning to prevent them. It also amplifies a full picture suspect activities around the time of the crime and make predictions about possible crimes. Additionally, it identifies performance opportunity with greater degree of accuracy for officers, investigators and intelligence officials in their agencies. Further, the impact properly predicts terrorism, crime and other threats to public security and safety with actionable intelligence. It is noteworthy that big data analytics drastically regulates the root causes of crime. Big data analytics have the potential to react timely and use accurate intelligence in order to transform every sphere of life and shaping the policy and laws to protect the core values of public.

3.4.5 Summary on Public Security and Safety

The study stressed that big data offers many research paths in protecting personal privacy, civil rights, catch criminals, scrutiny on particular individuals and making policies ensure fairness and prevent discrimination as presented in figure 8.



Figure 8. Outcomes of Big Data in Public Security and Safety

3.5 Personal Location Tracking

Personal location data is used to prosper new businesses and innovative business models that are reaching the lives of people across the globe. The personal location data is not confined to a single domain but it includes telecom, retail industry and media. This domain put forwards the potential value creation over the next decade to all stakeholders. To do this, this domain creates research challenges in building capable infrastructure, inventing innovative technology and professional capacity building. In addition, the preparation of qualified high volume of personal location data opens many more research paths for present big data researchers.

3.5.1 Big Data Resources in Personal Location Tracking

The implementation of GPS and GPRS technologies in handheld devices has turn into mountains of personal location data rapidly. Moreover, the enormous amount of personal location data stored in primary resources such as cell tower triangulation data of mobile devices, POS terminal data of personal cards and GPS chips data of navigation systems. In addition, a combination of navigation devices, cell tower tracking and smart phones also accumulate personal location data exponentially.

The personal location data growing rapidly as navigational devices updates data frequently, cell towers generates many users data and smart phones accumulates through fast growing sources. Also, the data generated by all these different types and sources of devices is in different formats. To conduct the analysis, personal location tracking domain on such data poses a significant series of practical and research challenges.

3.5.2 Big Data Technologies in Personal Location Tracking

The big data technology in personal location data domain aim to collect, aggregate, analyze, interpret, create, communicate and visualize the potential value effectively. The present technologies used in this domain related to smart routing, automotive telematics and mobile phone location based services are limited to deliver the expected services to all stakeholders.

All of these aspects after evolving personal location data domain pose significant challenges on ongoing innovation in technologies and techniques such as smart routing, path intelligence, shopping patterns, historical buying patterns, customer demographics, machine learning technologies and collecting & analysing data are future promising areas of research.

3.5.3 Big Data Applications in Personal Location Tracking

Big data analytics applied in this domain in three major categories on personal location data. Location-based applications and services for individuals is a category that includes smart routing, automotive telematics, and mobile-phone based location services. The second category is the organizational use of individual personal location data that includes geo-targeted advertising, electronic toll collection, insurance pricing, and emergency response. Third is the macro-level use of aggregate location data that includes urban planning and retail business intelligence.

3.5.4 Big Data Impact in Personal Location Tracking

Applying personal location data has the potential to raise economic value to individual consumers and organizations over the next decade, in the process catalyzing the development of a wide range of innovative businesses across many sectors. Smart navigation applications alone may offer value to global consumers in time. In addition, geo-targeted advertising is emerging as a highly effective marketing vehicle that could represent global advertising. Executives and policy makers need to work together to enable the growth of this data domain and unleash its full potential.

3.5.5 Summary on Personal Location Tracking

Applying personal location data has the potential to provide innovative businesses across many domains. Smart navigation applications and Geo-targeted advertising are working together to enable the growth of this domain and unleash its full potential outcomes are listed in figure 9.



Figure 9. Outcomes of Big Data in Personal Location Tracking

4. CONCLUSIONS

The authors in the present paper give a general overview of big data analytics supported by technological advancements which inspired to take up the comprehensive study. This study further highlights the evaluation and characteristics of big data. Later, literature survey is presented which is again confined to the basic idea behind the present study.

The proposed study mainly focussed on five key domains namely health care, retail industry, public governance, public security and safety and personal location tracking to explore promising research opportunities along with challenges in all possible areas for creating tremendous untapped potential value. In each domain, the study initially concentrates on the sources of big data and its potential hidden insights. Subsequently, it highlights the ongoing big data technologies and techniques to along with associated challenges. Further, it presents the possible productive big data application to improve efficiency and effectiveness to produce actionable intelligence. Afterwards, it pays the attention on the big impact of big data to capture value addition in each service. Finally, it represents key observations and outcomes of the respective domain in the lines of present research study.

Key observation of present study:

- Make use of big data easily available to appropriate stakeholders in a timely way that generates great value. Certainly, this aspect of creating value is a prerequisite for all other performance parameters in creating data transparency for all domains to capture that potential.
- Big data enables experimentation to discover customer centric needs, interpret variability in tremendous amount of data and improve performance of the system comprehensively.
- In order to provide target services and to meet individual needs, big data segment population through customized actions.
- Sophisticated big data analytics substantially improve human decision making all the time, minimize risks and bring to light valuable insights with automated algorithms.
- Big data facilitates all enterprises to create innovative products and services, improve existing ones and invent entirely new business models.
- In addition in the lines of present study, several key issues and organizational boundaries to be addressed such as data policies, technology & techniques, organizational change & talent, access to data and industry structure have been identified.

The present study systematically presents all possible potential future research challenges and opportunities to the present researchers at each stage of big data analytics across all applicable key domains.

ACKNOWLEDGEMENTS

The authors recorded their acknowledgements to Shri Vishnu Engineering College for Women (Autonomous), Bhimavaram and Andhra University, Visakhapatnam authorities for their constant support and cooperation.

REFERENCES

- [1] Amir Gandomi, Murtaza Haider, "Beyond the hype: Big data concepts, methods, and analytics", International Journal of Information Management, pp: 137-144, 2015.
- [2] Ashwin Belle, Raghuram Thiagarajan, "Review Article - Big Data Analytics in Healthcare", Hindawi Publishing Corporation, pp: 1-17, 2015.
- [3] J. Archana, E.A.Mary Anita, "A Survey Of Big Data Analytics in Healthcare and Government", Procedia Computer Science, Elsevier, pp:408-413, 2015.
- [4] V. Morabito, "Big Data and Analytics for Government Innovation", Springer International Publishing Switzerland, pp: 23-45, 2015.
- [5] Bellagio Big Data Workshop Participants, "Big data and positive social change in the developing world: A white paper for practitioners and researchers", Oxford Internet Institute, pp: 1-35, 2014.
- [6] Chuck Ballard, Cindy Compert, "Information Governance Principles and Practices for a Big Data Landscape", IBM, pp: 1-280, 2014.

- [7] Gang-Hoon Kim, Silvana Trimi, "Big-Data Applications in the Government Sector", communications of the ACM, vol. 57, No. 3, pp: 77-85, 2014.
- [8] Hewlett-Packard Development Company, "Business White Paper – Big Data and Health Care", pp: 1-8, 2014.
- [9] Intel, "Getting Started with Big Data Analytics in Retail", Solution Blueprint, pp: 1-13, 2014.
- [10] Kathryn Howe, "Beyond Big Data: How Next-Generation Shopper Analytics and the Internet of Everything Transform the Retail Business", Cisco, pp: 1-10, 2014.
- [11] The Parliamentary office of science and technology, "Big Data, Crime and Security", pp: 1-5, 2014.
- [12] Wullianallur Raghupathi, Viju Raghupathi, "Big data analytics in healthcare: promise and Potential", Raghupathi and Raghupathi Health Information Science and Systems, pp: 1-10, 2014.
- [13] Center for Media Justice, "Consumers, Big Data, and Online Tracking in the Retail Industry", A Case Study of Walmart, pp: 1-30, 2013.
- [14] Jean Yan, " Big Data, Bigger Opportunities", President Management Council Inter-agency Rotation Program, pp: 1-27, 2013.
- [15] Jimeng Sun, Chandan K. Reddy, "Big Data Analytics for Health Care", SIAM International Conference on Data Mining, Austin, pp: 1-112, 2013.
- [16] Joe Mullich, "Closing the Big Data Gap in Public Sector", Survey Report Real-Time Enterprise, pp: 1-4, 2013.
- [17] SaS Institute, " Big data for public safety", pp: 1-2, 2013.
- [18] Serendio, " Big Data Management and Predictive Analytics as-a-service for the Retail Industry", Executive Summary, pp: 1-6, 2013.
- [19] StackIQ, "White paper - Boosting Retail Revenue and Efficiency with Big Data Analytics" pp: 1-19, 2013.
- [20] Lin Gu, Deze Zeng, et.al., "Cost Minimization for Big Data Processing in Geo-Distributed Data Centers", IEEE Transactions On Emerging Topics In Computing, pp: 314-323, 2013.
- [21] Tata Consultancy Services, "The Emerging Big Returns on Big Data", Global Trend Study, pp: 1-106, 2013.
- [22] Adelaide O'Brien, "The Impact of Big Data on Government", Iron Mountain, pp: 1-12, 2012.
- [23] Alexander Olesker, " White Paper: Big Data Solutions for Law Enforcement", CTOLabs, pp: 1-8, 2012.
- [24] Chris Yiu, "The Big Data Opportunity-Making government faster, smarter and more personal", pp: 1-36, 2012.
- [25] Cognizant, "Big data is the future of Health Care", pp:1-7, 2012.
- [26] Danyel Fisher, Rob DeLine, et.al., "Interactions with Big Data Analytics", pp: 1-10, 2012.
- [27] IBM Institute, "Analytics: The real-world use of big data in retail", IBM Global Business Services, Saïd Business School at the University of Oxford, pp: 1-16, 2012.
- [28] Infosys, "Big Data Spectrum", pp: 1-61, 2012.
- [29] Srinivasan N, Rajeev Nayar, "Harnessing the power of big data, big opportunity for retailers to win customers", pp: 1-8, 2012.
- [30] Brad Brown, Michael Chui, and James Manyika, "Are you ready for the era of 'big data'?", McKinsey Global Institute, pp: 1-12, 2011.
- [31] McKinsey Global Institute, "Big data: The next frontier for innovation, competition, and productivity", pp: 1-156, 2011.
- [32] SaS Institute, "Big Data Meets Big Data Analytics", White Paper, pp: 1-13, 2011.
- [33] Sherif Sakr, Anna Liu, Daniel M. Batista, and Mohammad Alomari, "A Survey of Large Scale Data Management Approaches in Cloud Environments", IEEE Communications Surveys & Tutorials, Vol. 13, No. 3, pp: 311-336, 2011.
- [34] Steve LaValle, Eric Lesser, et.al., "Big data, analytics and the path from insights", MIT Sloan Management Review, pp: 1-13, 2011.
- [35] David Bollier, Rapporteur, "The Promise and Peril of Big Data", Aspen Institute, Washington, pp:1-66, 2010.
- [36] Kelton Research, "Global Survey: The Business Impact of Big Data" " , Avande, pp:1-4, 2010.

Authors

Dr. V.V.R. Maheswara Rao received his doctoral degree in Computer Science & Engineering at Acharya Nagarjuna University, Guntur, India. He is working as Professor in the Dept of Computer Science and Engineering Department at Shri Vishnu Engineering College for Women (Autonomous), Bhimavaram, AP, India. His Research interests include Data Mining, Web Mining, Big Data Analytics, Artificial Intelligence and Genetic Algorithms.



Dr. V. Valli Kumari holds a Ph.D. degree in Computer Science and Systems Engineering from Andhra University Engineering College, Visakhapatnam and is presently working as Professor in the same department. Her research interests include Security and privacy issues in Data Engineering, Network Security and E-Commerce. She is a member of IEEE and ACM.



N Silpa completed her M. Tech. degree in Computer Science and Engineering at Jawaharlal Nehru Technological University, Kakinada, India. She is working as Asst. Professor in the Dept of Computer Science and Engineering Department at Shri Vishnu Engineering College for Women (Autonomous), Bhimavaram, AP, India. Her Research interests include Data Mining, Web Mining, Big Data Analytics, Text Mining and Genetic Algorithms.