

# A Survey on Issues and Challenges of Big Data

Kavitha.G\* and Elango N.M\*\*

\*School of information Technology and Engineering  
kavitha.g@vit.a.c.in

\*\*school of information Technology and Engineering  
elango.nm@vit.ac.in

**Abstract:** In this era of information explosion, there has been a continuous growth in amount of data which are generally generated via social media, Internet of Things, may be in either structured or unstructured format which are referred as term “big data”. The most common challenges and opportunities arising is storage of data, developing an application for analyzing big data efficiently, and to design an appropriate mining algorithm to extract knowledge from big data. The aim of this paper is to give a brief introduction about the evolution and characteristics of big data followed with major issues and challenges of big data with some of the domain-specific areas such as cloud computing, Internet of Things, deep learning and data mining.

**Keywords:** Big data, cloud computing, data mining, Internet of Things, deep learning.

## Introduction

Data growth is one of the most important factors considered for several operations across the enterprises. Indeed; the top level management in organizations will analyze their business intelligence operations and also interact with individuals about the performance, outcome of the end product developed by them. The data which generated as a by-product of other activities such as Social media sites, smartphones, and other consumer devices including PCs and laptops have allowed billions of individuals around the world to contribute to the amount of data available and there by generating a tremendous amount of digital “exhaust data”. The continuous proliferations of data captured by organizations, the data creation is referred as big data, emerging as a widely recognized trend [1]. Also nearly about 2.5 quintillion bytes of data are created and around 90 percent of world’s data was created in the past two years alone. It also indicated that the volume of the business data will grow significantly every year shown in Figure 1.

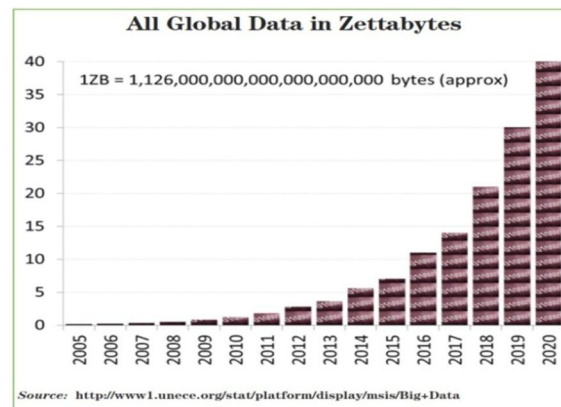


Figure 1: Growing of data volume [source: Graphic showing real and predicted growth of unstructured data generated between 2005-2020. Image courtesy of IDC][2].

The report generated by IDC expects the Big Data technology and services market to grow from \$6 billion in 2011 to \$32 billion in 2017, representing a compound annual growth rate (CAGR) of 31.7% or about seven times that of the overall information and communication technology (ICT) market [2]. The reports of [3] and [4] also stated that marketing of big data will be \$46.34 billion and \$114 billion by 2018, respectively. Though the marketing values of the above reports are different but these projections usually indicate that the scope of big data will grow rapidly in the forthcoming future [5]. The paper is organized as follows: section 2 focuses on the context and evolution of Big Data; section 3 it’s opportunities and challenges with respect to different information, communication, technologies, and the final section draws the conclusion and future work.

### The Context and Evolution of Big Data

Big data is a term generally referred to “ the continuous increase in volume of data which is difficult to store, process and analyze through current database technologies”[6]. Though the term “big data” referred as a large amount of data which is beyond the technologies capability to store, manage and processing efficiently”. Similarly [7] and [8] characterized the big data by three vs.: volume, variety, and velocity as was originally stated by Gartner and he defined the term big data into 3 parts of definition such as “high-volume, high-velocity, and high-variety information assets that demands cost-effective, innovative forms of information processing for enhanced insight and decision making.” [9]. IDC also defined big data technologies as “a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high velocity capture, discovery, and/or analysis.” [10]. As companies are more interested to gain maximum benefits from big data and analytics in turn producing prediction of future trends and also computing resources evolved to better handle data size and complexity, the evolution of big data are represented in figure 2.

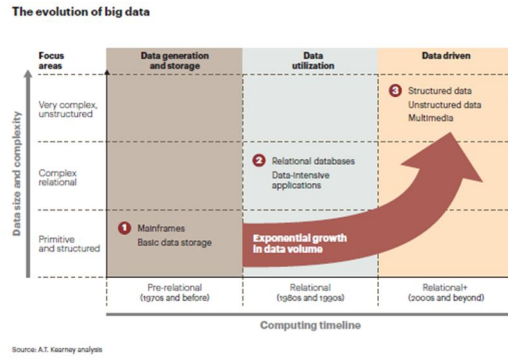


Figure 2: evolution of big data [source: A.T Kearney analysis] [11]

### Characteristics of Big Data

The characteristics of big data are usually defined based on HACE Theorem, as large-volume, Heterogeneous, Autonomous. The following characteristics of big data are proposed based on the above definitions as discussed in the previous section. Big data requires the integration of new technical architectures with the use of several analytical methods and tools to explore the new sources of business value. In general big data has been characterized are listed in Table 1.

Table 1: 5 v’s of Big data

Attributes	Nature	Examples
Volume	Huge amount of data that either requires huge capacity to store or number of transactional records, being in large numbers[12]	The data warehouse generated by Wal-Mart’s which includes nearly about 2.5 petabytes of information [13]. Dell has been initiated in developing the database that included 1.5 million records[14]
Variety	The acquisition of data are from different variety of source and formats [12]	Tata Motors analyzed 4 million text messages every month which includes about the product complaints, service appointments, announcements about new models and also connected with customer satisfaction polling (Agarwal and Weill, 2012-13).
Velocity	The instance of data generation and data delivery [12]	Amazon manages a constant flow of new products, suppliers, customers, and promotions without compromising promised delivery date and also the data generated by individual click stream from the web can be used for behavior analysis [14].
Veracity	To gain reliable prediction, data requires to be analyzed on it[15]	Data replication which was scattered throughout the various data marts in embaying which had been resolved through an internal website (data hub) which enables managers to filter data replication [15].
Value	The extent of worthy insights to generate economically worthy information	Premier Healthcare Alliance used enhanced data sharing and analytics to improve patient outcomes while reducing spending by US\$2.85 billion[16]

## Classifications of Big Data

According to [6-17] it has been categorized into 5 types which has been listed in Table 2.

Table 2: Different Types of big data

Types of data	Sources
Web and social media data	Clickstream and Interaction data from Facebook, Twitter, LinkedIn, blogs [17].
Machine to machine data	Readings from remote Sensors, system log files and other sign devices [17].
Big transaction data	All kind of transactional data which are in form of both structured and semi structured [17].
Biometric data:	Finger prints, genetics, handwriting, retinal scans, x-ray and other medical images, blood pressure, pulse and pulse-oximetry readings, another similar types of data [17].
Human-generated data:	Unstructured and semi-structured data such as EMR's, physician's notes, email[17].

## Related Technologies

### Cloud Computing

Cloud computing is one of the emerging technologies, establishing itself as the next generation of IT industry and business. It has become a powerful architecture in order to perform a computational tasks (store, process and analyze) for complex large-scale databases such as big data [18]. Therefore, the indent of cloud computing is not only to provide computational and processing tasks for big data, but also act as a service mode. It has also begun to integrate different architectural frameworks to provide services in parallel data processing environment which helps the data analyzers to access the various could services such as Peas, Iaas, and Saas to deploy their programs (application) [19]. Due to limited processing capability, storage capacity, and battery life time of each device [20]. However, the limitation is all about the computational and data storage restrictions [21]. Though it is widely accepted that cloud computing is mass technology which resolves the above limitations but we are faced with new challenges from the applications which are used by the organizations. Some of the key research issues are listed in the table below:

Table 3: Issue on cloud computing with big data

Research Challenges	Issues
Scalable	Storage of data in distributed [22]
Availability	Accessible on demand of data stored in cloud[23]
Heterogeneity	Incompatible format of data representation[24]
Preserving	How to preserve privacy of intermediate datasets[25]
Security	Encapsulating sensitive data with secure key management in access to big data [26]

### Big Data and IOT

The data generated for big data emerges from a variety of sources in very large amounts and often in real-time settings with the encapsulation of cheap sensors, numerous and various apps on mobile devices, and an increasingly connected world that sometimes does not even require human intervention. The different kinds of data generated from different fields such as environmental, geographical, astronomical, logistics, mobile equipment, transportation facilities, and home appliances all of which will be data acquisition cores which introduces the technology such as Internet of Things [IOT]. The origin of data for IOT is generated from Radio Frequency Identification (RFID), other tracking and sensor devices. It is stated as “an increasingly connected world that sometimes does not even require human intervention (as shown in the increasing development of the Internet of Things [IoT])” [27].

Intel identified and pointed out that they are about three basic features to be confirmed with big data paradigm [28]:

- (i) Sensor based terminals that generating masses of data.
- (ii) Most probably there will be semi-structured or unstructured data format.
- (iii) The data generated by IOT will be more precise only when it is analyzed.

Min Chen et al [47-29] stated that the effect of IoT in real time will be get into great success only when there is an effective integration of big data and cloud computing.

Recently the application of IOT such as Smart city is a hot research area based on the application. It helps to obtain better information support in decision making for managing resources. For instance, the department of Park Management of Dade

County saved one million USD in water bills due to timely identifying and fixing water pipes that were running and leaking this year.

*Research Issues*

Based on our IOT adoption, usage, and impact the potential future research directions are tabulated in tabular format. It is delivered into two level of dimension such as individual and organization level.

*Individual level:* There is a continuous increased hesitancy among individuals on adopting location specific, tracking using sensor based devices due to privacy and security concerns. Several important adoption issues related with this technology at the individual level [30] is listed in Table 2:

- How will usage of the IoT be influenced by views and usage of personal data?
- What human behavioral changes are likely to occur due to pervasive diffusion of smart devices and smart products connected to the IoT?

*Organizational level:* In most of the organization they are almost enabled to develop their products based on RFID-enabled IoT applications. One of the most important perspective is unveiling the organizational readiness in relation to the business value realized from IoT-enabled applications [32]. Some of the research challenges are listed as below:

- What is the role of big data analytics in process re design, efficiency benefits, effectiveness, benefits, and strategic benefits?
- Evaluation on applicability of context-aware computing techniques towards IoT.
- Insecure Web Interface.
- Insecure Network Services.
- Lack of Transport Encryption.
- Insufficient Security Configurability.

*Deep Learning*

In the growing digital world the two hottest trends, deep learning and Big Data, are growing rapidly. According to the report generated by National Security Agency stated that the amount of digital information acquired has grown by about 1826 petabytes, nine times in volume around five years, and by 2020 it will reach 35 trillion gigabytes [32]. Due to the unprecedented increase in size of non-traditional data generated in the growing digital world and there being an increasing demand in development of various machine learning techniques which is highly utilized for knowledge discovery [33]. According to Y. Bengio and S. Bengio who defined “deep learning as a machine learning technique that use supervised and/or unsupervised strategies to automatically learn hierarchal representations in deep architectures for classification” [34]. The role of big data in deep learning by providing predictive analytic solutions with improved processing power and the advances in graphic processors. Some of the deep learning methods such as DBN, CNN, Disbelief, Sparse auto encoder, COTS HPC were used for analyzing big data [35].

Table 4: Security challenges with respect to deep learning with big data

Different aspects of Big data	Security challenges.
Infrastructure Security	Secure Distributed Processing of Data and Non-Relational Data-Bases[37]
Data Privacy	Cryptographic solutions for Data security[38] Granular Access Control[39]
Data Management and Integrity	Secure Data Storage and Transaction Logs[40] Data Provenance[41]
Reactive security	Filtering and Validation[42] Supervising the Security Level in Real-Time[43]

*Research Challenges*

Though the current deep learning methods are applied in big modalities. The common issues arising are:

- Is the system performance significantly accessed for enlarged modalities?
- Are deep learning architectures appropriate for feature fusion with heterogeneous data?

**Security and Privacy Issues in Big Data**

One of the most valuable assets of an organization is the data which is generated by the users ranging from scientific advances to business intelligence. Thus the organizations are much more concentrated on security, and preservation of privacy of their outsourced data needs to be tackled. Some of the well-defined security aspects such as firewalls and demilitarized zones (DMZs) are said to be no more effective for big data. The secure services can be developed only with

respect to heterogeneous infrastructure As most of the social network users shares their personal data and also the data generated by the mobile devices and computers as when they accessed by the users. The loosing data and content control are the two major issues are highly impact over on privacy [36].The objective of creating security based solutions is to preserve its confidentiality, integrity and availability.

### Data mining and big data

Due to emergence of fast and massive streams of data, analyzing on those will lead to new valuable insights and theoretical concepts [41-44] Extraction of hidden information and insights from big data done through data mining. Even though the process of knowledge discovering are done by applying existing data mining algorithms and techniques from conventional databases yet it is tends to be challenging task on extracting the information from the three Vs of big data [45].There are about two major factors will be considered on knowledge discovery process such as data access time and efficiency of the mining algorithms[46,47].The need for designing and implementing very-large-scale parallel machine learning and data mining algorithms has remarkably increased.

#### Major challenges and issues of big data mining

The major challenge is that mining hidden patterns from big data which are in heterogeneous type and also designing multi-model systems for complex type of heterogeneous multimedia data is said to be another major challenge related big data mining. Some of the major challenged related to scalability, accuracy, redundancy are listed in the below Table 5.

Table 5: issues and challenges of big data mining

Major Issues	Challenges
Redundancy	Design of efficient indexing schemes [48].
Scalability	Design of parallel processing architectures[49]
Accuracy	Data validation and provenance tracing [66-50].
Garbage Mining	sustainability of our digital environment by mining the garbage values (outdated, corrupted, and useless)[51]

### Conclusion

There will be continuous process on extracting useful amount of information and knowledge from data that are generated by users who come up with new ways to message and process data. This paper initiates in identifying major issues and challenges with different research domains with respect to big data. Future review will concentrate on detailed description about issues and also to explore their existing solutions associated with different domain allied with big data.

### References

- [1] Chen, M., Liu, Y., & Mao, S., "Big Data: A Survey", MONET, 19,2014, 171-209.
- [2] Worldwide Big Data Technology and Services 2012–2015 Forecast (IDC #233485, March 2012)
- [3] Taft DK. Big data market to reach \$46.34 billion by 2018, EWEEK, Tech. Rep. 2013.[Online].Available:<http://www.eweek.com/database/big-data-market-to-reach-46.34-billion-by-2018.html>.
- [4] Research A. Big data spending to reach \$114 billion in 2018; look for machine learning to drive analytics, ABI Research, Tech. Rep. 2013. [Online].Available:<https://www.abiresearch.com/press/big-data-spending-to-reach-114-billion-in-2018-100>.
- [5] Tsai et al., "Big data analytics: a survey", Journal of Big Data, 2:21.2015 DOI 10.1186/s40537-015-0030-3
- [6] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A.H. Byers, "Big data: The next frontier for innovation, competition, and productivity", 2011.
- [7] P.Zikopoulos,K.Parasuraman,T.Deutsch,J.Giles,D.Corrigan,Harness , "The Power of Big Data", The IBM Big Data Platform, McGraw Hill Professional,2012.
- [8] J.J.Berman, "Introduction: Principles of Big Data", 2013,xix–xxvi.
- [9] Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s, By Svetlana Secular, Gartner, Inc. 27 March 2013. [online] [http://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs./](http://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/)
- [10] It's Much More Than Technology: Big Data and Analytics Adoption Requires Focus on Strategy, Skills, and Process (IDC #cUS23693312, September 2012)
- [11] "Big data and Creative Destruction of Today's Business Models", AT Kearney ,2013. [Online] Available: [https://www.atkearney.com/strategic-it/ideasinsights/article//asset\\_publisher/LCcgOeS4t85g/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192](https://www.atkearney.com/strategic-it/ideasinsights/article//asset_publisher/LCcgOeS4t85g/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192)
- [12] models/10192
- [13] Russo, P. "Big Data Analytics," TDWI Best Practices Report, Fourth Quarter,2011.
- [14] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. "Big Data: The Next Frontier for Innovation, Competition, and Productivity," McKinseyGlobalInstitute,2012. [online] Available

- ([http://www.mckinsey.com/insights/mgi/research/technology\\_and\\_innovation/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation); accessed August 4, 2012).
- [15] Davenport, T. H. 2006. "Competing on Analytics," *Harvard Business Review* (84:1), p. 98-107.
- [16] Bloomberg Business week. 2011. "The Current State of Business Analytics: Where Do We Go from Here?" Bloomberg Business-week Research Services ([http://www.sas.com/resources/asset/busanalyticsstudy\\_wp\\_08232011.pdf](http://www.sas.com/resources/asset/busanalyticsstudy_wp_08232011.pdf)).
- [17] Deutsch, T. "Recapping 2014: Significant Trends for Big Data" *IBM Data Magazine*, March 28, 2014.
- [18] IBM. 2011. "The 2011 IBM Tech Trends Report: The Clouds are Rolling In...Is Your Business Ready?," [online] Available (<http://www.ibm.com/developerworks/techtrendsreport>; accessed August 4, 2012).
- [19] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers "Big data: The next frontier for innovation, competition, and productivity", *McKinsey Global Institute*, May 2011
- [20] L.Huan, "Bigdata drives cloud adoption in enterprise", *IEEE Internet Computing*, 2013, Volume:17 Issue:4, 68–71.
- [21] D.Warneke, O.Kao, Nephel: efficient parallel data processing in the cloud, in: *Proceedings of the 2nd workshop on many-task computing on grids and supercomputers*, ACM, 2009, p.8.
- [22] Khan, AbdulNasir, et al., "BSS:block-based sharing scheme for secure data storage services in mobile cloud environment", *The Journal of Supercomputing*, 2014, 1–31.
- [23] D. Zisis and D. Lekkas, "Addressing Cloud Computing Security Issues", *Future Generation Computer Systems*, Vol. 28, No. 3, 2012, pp. 583-592. [online] Available: <http://www.sciencedirect.com/science/article/pii/S0167739X10002554>.
- [24] D.Che, M.Safran, Z.Peng, From big data to big data mining: challenges, issues, and opportunities, in: B.Hong, X.Meng, L.Chen, W.Winiwarter, W.Song (Eds.), "Database Systems for Advanced Applications", Springer, Berlin Heidelberg, 2013, pp. 1–15.
- [25] Z.Xuyun, L.Chang, S.Nepal, S.Pandey, C.Jinjun, A Privacy Leakage Upper Bound Constraint-Based Approach for Cost-Effective Privacy Preserving of Intermediate Datasets in Cloud, Parallel and Distributed Systems, *IEEE Transactions*, 2013, pp. 1192–1202.
- [26] Z.Sanaei, S.Abolfazli, A.Gani, R.Buyya, "Heterogeneity in mobile cloud computing: taxonomy and open challenges", *IEEE Commun. Surv. Tutor.* (2013) 1–24.
- [27] P. Mell, T. Grance, "The NIST definition of cloud computing (draft)", *NIST Spec. Publ.* 800, 2011.
- [28] T.Gunarathne, B.Zhang, T.-L.Wu, J.Qiu, "Scalable parallel computing on clouds using Twister 4 Azure iterative Map Reduce", *Future Generation. Computer Systems*, vol 29, 2013, pp. 1035–1048.
- [29] A. Driscoll, J.Daugelaite, R.D.Sleator, " 'Big data', Hadoop and cloud computing in genomics", *J.Biomed.Inform.* vol 46, 2013, pp. 774–781.
- [30] J.Dean, S.Ghemawat, "MapReduce: simplified data processing on large clusters, Communication", *ACM*, vol 51, 2008, pp. 107–113.
- [32] D.Bollier, C.Firestone, M., "The Promise and Peril of BigData", *Aspen Institute, Communications and Society Program* Washington, DC, USA, 2010.
- [33] Dr Gloria González Fuster and Dr Amandine Scherrer, Directorate general for internal policies policy department c: citizens' rights and constitutional affairs, "big data and smart devices and their impact on privacy, Study for the LIBE Committee", 2015.
- [34] Intel IT Center, Peer Research Big Data Analytics Intel's IT Manager Survey on How Organizations Are Using Big Data. August 2012.
- [35] Min Chen · Shiwen Mao · Yunhao Liu " Big Data: A Survey", *Mobile Netw* April (2014) 19:171–209] DOI 10.1007/s11036-013-0489-0
- [36] Rubinstein, I.S., "Big Data: The End of Privacy or a New Beginning?", *International Data Privacy Law*, 3(2), 2013, pp. 74-87.
- [37] Hsinchun, C., Chiang, R.H.L., and V.C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", *MIS Quarterly*, 36(4), 2012, pp. 1165-1188.
- [38] National Security Agency. The National Security Agency: Missions, Authorities, Oversight and Partnerships [Online]. Available: [http://www.nsa.gov/public\\_info/\\_les/speeches\\_testimonies/2013\\_08\\_09\\_the\\_nsa\\_story.pdf](http://www.nsa.gov/public_info/_les/speeches_testimonies/2013_08_09_the_nsa_story.pdf)
- [39] J. Lin and A. Kolcz, "Large-scale machine learning at twitter," in *Proc. ACM SIGMOD*, Scottsdale, Arizona, USA, 2012, pp. 793\_804.
- [40] Y. Bengio and S. Bengio, "Modeling high-dimensional discrete data with multi-layer neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 12, 2000, pp. 400\_406
- [41] X.-W. Chen and X. Lin: "Big Data Deep Learning: Challenges and Perspectives", *VOLUME 2*, 2014, pp 514-525.
- [42] Shrivastva, K, 2014. " Big Data Privacy Based On Differential Privacy a Hope for Big Data". *Sixth International Conference on Computational Intelligence and Communication Networks*, vol 1, 2014, 777-778.
- [43] Boja C, Pocovnicu A, Batagan L. Distributed Parallel Architecture for Big Data[J]. *Informatica Economica*, 2012, 16(2): 116-127
- [44] Advantech. (2013). Enhancing Big Data Security. Retrieved from [http://www.advantech.com.tw/nc/newsletter/whitepaper/big\\_data/big\\_data.pdf](http://www.advantech.com.tw/nc/newsletter/whitepaper/big_data/big_data.pdf)
- [45] Cloud Security Alliance. (2013). Expanded Top Ten Security and Privacy Challenges. Retrieved From [https://downloads.cloudsecurityalliance.org/initiatives/bdwwg/Expanded\\_Top\\_Ten\\_Big\\_Security\\_and\\_Privacy\\_Challenges.pdf](https://downloads.cloudsecurityalliance.org/initiatives/bdwwg/Expanded_Top_Ten_Big_Security_and_Privacy_Challenges.pdf)
- [46] Mr. Yannam Apparao, Mrs. Kadiyala Laxminarayamma, "Security Issue on "Secure Data Storage and Transaction Logs" In Big Data, *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)* ISSN: 2347-5552, Volume-3, Issue-3, May 2015
- [47] Wang-Chiew Tan, "Research Problems in Data Provenance", *IEEE Computer Society Technical Committee on Data Engineering* Available online: <https://pdfs.semanticscholar.org/de74/d5827b9ca8a429da2d9a1242dd699ec51700.pdf> Accessed on: 24<sup>th</sup> March 2016.

- [49] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public auditability and data dynamics for storage security in cloud computing," *Parallel and Distributed Systems, IEEE Transactions on*, 22(5):847–859, May 2011
- [50] J. Onieva, J. Lopez, and J. Zhou, "Secure Multi-Party Non-repudiation Protocols and Applications," *Advances in Information Security Series*, Springer, 2009.
- [51] Berkovich, S., Liao, D.: On Clusterization of big data Streams. In: 3rd International Conference on Computing for Geospatial Research and Applications, article no. 26. ACM Press, New York (2012)
- [52] Beyer, M.A., Laney, D.: The Importance of 'Big Data': A Definition. Gartner (2012)
- [53] Zhang, X., Ai, J., Wang, Z., Lu, J., Meng, X.: An Efficient Multi-dimensional Index for Cloud Data Management. In: 1st International Workshop on Cloud Data Management, pp.17–24. ACM Press, Hong Kong (2009)
- [54] Borkar, V.R., Carey, M.J., Li, C.: big data Platforms: What's Next? *ACM Crossroads* 19(1), 44–49 (2012)
- [55] Yin, X., Han, J., Yu, P.S.: Truth Discovery with Multiple Conflicting Information Provider on the Web. In: 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Jose, California, pp. 1048–1052 (2007)
- [56] B R Prakash and Dr. M. Hanumanthappa , Issues and Challenges in the Era of Big Data Mining, *International Journal of Emerging Trends & Technology in Computer Science*, ISSN 2278-6856, Volume 3, Issue 4 July-August 2014
- [57] Agrawal, D., Bernstein, P., Bertino, E., et al.: Challenges and Opportunities With big data– A Community White Paper Developed by Leading Researchers Across the United States (2012), <http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>
- [58] Stephen Kaisler et al., Big Data: Issues and Challenges Moving Forward, 46th Hawaii International Conference on System Sciences (2013), IEEE DOI 10.1109/HICSS.2013.645