

# An Analytical Research on the Big Data Tools and Techniques

<sup>1</sup>Yoshitaka Akiyama and <sup>2</sup>Aravindhnan K

<sup>1</sup> Department of Computer Science and Engineering, Tokyo Institute of Technology, Japan.

<sup>2</sup>Department of Computer Science and Engineering, SNS College of Engineering, Tamil Nadu, India.

<sup>1</sup> yoshitakaaki@gsic.titech.ac.jp, <sup>2</sup> aravindhnan03@gmail.com

## ArticleInfo

International Journal of Advanced Information and Communication Technology

([https://www.ijaict.com/journals/ijaict/ijaict\\_home.html](https://www.ijaict.com/journals/ijaict/ijaict_home.html))

<https://doi.org/10.46532/ijaict-2020026>

Received 14 April 2020; Revised form 18 May 2020; Accepted 30 July 2020; Available online 05 August 2020.

©2020 The Authors. Published by IJAICT India Publications.

This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Abstract**—A massive segment of information terabytes is produced daily due to the advent of technological systems and modern systems such as cloud computing and the Internet of Things (IoT). The evaluation of the big data analytics impacts significant effort of the many dimensions considered in the process of extracting data that is essential for the process of making proper decisions in an organization. In that regard, big data analytics is the immediate area of concern for research and development experts. The primary purpose of this paper is to provide an analysis of the big data analytics, including the issues and tools linked to the process of extensive information analysis. Moreover, this academic paper presents an avenue that stimulates the exploration of big data at an incredible measure. As such, this research provides a new platform of analytics to formulate a remedy to the problems and the gaps available in research.

**Keywords**— Big Data, Internet of Things (IoT), Social Network Analysis (SNA), Machine Learning (ML)

## 1. Introduction

In the modern age, the production of data is enhanced by the involvement of multiple sources which incredibly facilitates the rapid transition from modern technological aspects. As a result, these transformations have led to the advancement of massive data. This significantly provides a transitional breakthrough in the various fields which are essential for the process of collecting massive datasets. Generally, it denotes the gathering of complicated and extensive sets of data that are difficult to analyse and process with the application of ancient database processing tools or information management tools. The present unstructured, structured and semi-structured data formats are in petabytes and beyond. The arrangements are considered from 3Vs to about 4Vs which represent velocity, variety and volume. The aspect of size denotes the massive amount of information that is produced daily, whereby speed represents the rate at which knowledge grows and how this information is gathered for analysis. The variety on the other side recommends the data about types such as unstructured, semi-structured and structured.

The critical purpose of the massive data analysis is to enhance the processing of data characterized by high-

volume, veracity, velocity and variety based on the application of the computational and ancient intelligent methods. Analytics Haider and Gandomi analyse some of the extraction methodologies used in the process of retrieving valuable data. Nonetheless, the precise definition of massive data is not provided which makes it appealing to believe that big data might be problem specific. As such, this incredibly helps researchers to come up with more fundamental decisions with enhanced optimization and insight discovery, whereas being more cost-effective and tech-savvy. It is projected that the advancement of massive data is about twenty-five billion which is significantly advancing as time goes. Based on the standby of data and information transfer technologies, enormous data is a significant and impetus that possesses substantial application for the upcoming generation of technological organizations. These perspectives are based on the third platform which significantly refers to the massive data, Internet of Things (IoT), social organizations and cloud computing.

Information warehouses have significantly been involved in the management of massive sets of data. As such, facilitating the extraction of the precise skillsets from the present massive data sets is a significant concern. A significant portion of the available approaches in the information mining sectors is not reliable enough to deal with the massive sets of data. The major concern in the evaluation of massive data is the absence of coordination that should be there between data systems and analytics tools such as statistical analytics and data mining prospects. These problems categorically come up when researchers purpose is to perform skillset representation and discovery processes for the enhancement of practical application. The major issue is how a quantitative description of the massive data characteristics will be enhanced. There is a necessity for the epistemology effect meant to evaluate the revolution aspect of datasets. Moreover, the research on the complexity theory around the massive data will significantly aid the process of comprehending the first formation and characteristics that define the complex dataset patterns, simplifying

representations and obtaining effective skillsets for the process of abstraction. Apart from that it is possible to obtain guidance in the designing of computational algorithms and models on massive data.

Significant research has been done by academic researchers on massive data and the trends it follows. Nonetheless, it is to be confirmed that all the information present in the format of massive data is unreliable for evaluating and enhancing the process of making precise decisions. The academia and companies are more concerned with the process of dissemination of big data findings. This research significantly deals with the problems of massive data and the existing tools and techniques, whereas stressing on the available research problems on massive data. To achieve the paper's rationale, this research begins with part 1 that analyses the problems of big data analytics and its fine-tuning. Part 2 includes a literature analysis of the available research problems which will aid in the processing of massive data to enhance the provision of significant skillsets. Part 3 describes the big data insights, techniques and tools while Part 4 provides the projected research aims. Part 5 concludes the paper.

## 2. Problems of Data Analytics and Its Fine Tuning

Over the past few decades massive data has been recorded in several domains such as retail, interdisciplinary scholastic researchers, healthcare, biochemistry and public administration domains. The web-centred application faces massive sets of information frequently which includes internet, documents or texts, social computing and the Internet Search Indexes (ISI). Social computing incorporates Social Network Analysis (SNA), IoT, the online communities, recommended internet frameworks, projected markets and reputational systems. Internet Search Indexes (ISI) considers Scopus, IEEE Explorer and Thomson's Reuters among others [1]. About the merits of massive data users can potentially become exposed to novel opportunities in the skillset management of tasks meant to complete future researches. Nonetheless, these opportunities typically consider some problems.

To effectively mitigate these concerns, it is essential to get acquainted with the essential computing complexities, data securities and computational methodologies in the process of effectively analysing the big data in 2020 [2]. For instance, an incredible portion of statistical methodologies performing incredibly for small information sizes which do not have to accumulate to datasets of large volumes. In the same case, a lot of computational methods performing well for minor information sets encounter some fundamental problems that are essential in our process of analysing big data. Many problems which the healthcare industry faces are significantly being investigated by a lot of researchers [3].

In this paper, the problems of massive data analytics have been categorized into four classes namely: information analysis and storage, skillset discovery and computing complexity, visualization and scalability of information and data security. The issues of big data have been discussed in the following mentioned subsections.

### 2.1 Information Analysis and Storage

Data analysis and storage have existed over the past few years. In recent years, the amount of information significantly has grown based on the application of various methods such as aerial sensory methods, mobile appliances, radiofrequency, remote sensing and identification determiners among other methodologies. These information sets are stored, about their expenses which might later be deleted or ignored in case space is unavailable for storage. In that case, the initial problem for massive data is storage spaces or complete input and output velocity [4]. Information accessibility has to be a major priority that enhances the process of skillset representation and discovery [5]. In essence, the major concern makes it relevant to conclude that data can be obtained promptly and easily to enhance the process of information analysis. Over the past few decades, data analytics makes use of multiple disk drives to facilitate the storage of information. However, this data house is considered slower in terms of random output and input performance compared to sequential input and output.

So, to effectively deal with this challenging concern this storage limitation has to be controlled by the phase-change memory (PCM) and the solid-state drive (SSD). Nonetheless, the present storage methods cannot include the necessary performance to manage big data. Apart from that the diversification of information is another key issue of massive data analytics. Considering the constantly advancing technology surrounding datasets, data mining tasks have incredibly developed [6]. Moreover, the selection and reduction of data are among the fundamental tasks that are inhibited in the process of dealing with massive sets of data. This provides an unprecedented problem for data analytics, whereby the present algorithms might potentially not act within the required timeframe when users are dealing with present extreme datasets. The automation of this process and the development of novel Machine Learning (ML) algorithms inhibit the process of ensuring consistency which possesses a major concern in recent years.

To add to the clustering of massive sets of data which are essential in the process of analysing massive data is something researchers should consider. The recent advancements like MapReduce and Hadoop make it appealing for experts to gather massive amounts of both unstructured and structured data during a specific timeframe. The major engineering issue is how the dataset can be analysed in the process of ensuring better skillsets have been classified. The standard process is meant to end and transform the unstructured and semi-structured data process into structured before applying information mining algorithms meant to enhance the extraction of knowledge. Researchers [7] earlier introduced this framework to evaluate data. In the same case, detailed analyses of information evaluation of public comments were also analysed [8]. The key issue in this case is to consider more on the designing of storage frameworks and to push onto the incredible information evaluation tools which are meant to assure data output when data has been retrieved from various sources. With that being said the architecture of machine learning algorithms meant to evaluate data is

fundamental to enhance scalability and efficiency significantly.

### 2.2 Skillset Discovery and Computing Complexity

The discovery of skills and its representation is an incredible concern in massive data. This is because it contributes to the sub-fields which include preservation, authentication, data retrieval and management and data representation. With that regard, there are different tools for skillset discovering and computing representation like the fuzzy set, soft set, formal concept evaluation, major component analysis and the rough set [9]. Apart from that a lot of hybridized methods are incredibly formulated to evaluate actual-life issues. All these methods are dependent on the data issue. Furthermore, a number of the technique might be unavailable or unsuitable for massive datasets in a sequential PC system which incredibly poses another challenge. With that in mind several techniques compose incredible features of scalability and parallel computing. Due to the dimension of data-keeping on increasing incredibly, the existing tools might not be effective in processing the sets of data purposed to retrieve useful data.

The most recommended methodology in the process of managing datasets is making use of data marts and warehouses which are considerate of the information warehouses and facility evaluation [10]. The evaluation of massive sets of data necessitates significant computation complexities. The key concern here is to deal with the uncertainties and inconsistencies available in the datasets. Generally, the systematics models of computational complexities are utilized. As such, it might be difficult to consider the application of a mathematics framework which is incredibly applied to massive data. Nonetheless, domain-centred analytics can be executed easily through the process of comprehending these certain complexities. The main concern is to deal with the uncertainties and inconsistencies available in the datasets. Generally, the systematics model of computing complexities is applied hence making it less-appealing to establish a more comprehensive mathematical framework applied in massive data. Until now, a lot of survey and research has been conducted based on this problem direction using machine learning methods with applicable memory requirement [11]. The major objective of this analysis is to minimize computing cost complexities and processing. Nonetheless, present massive data analytics and tools are characterized by poor performance in the process of dealing with computing uncertainty, complexities and inconsistencies [12]. As a result, this leads to a significant issue to formulate technologies and methods meant to handle computational uncertainties, complexities and data inconsistencies in an incredible manner.

### 2.3 Visualization and Scalability of Information

Ensuring data security and scalability is a significant concern as a massive data analytics methodology. Over the past few decades, data researchers have considered focusing on how data will be accelerated and how its velocity will be boosted about Moore's law. It is vital to formulate sampling, multiple resolutions and online data analytics methods. The incremental methods are known to have incredible scalability competency in the avenue of

massive data evaluation. As the information dimension keeps on scaling significantly compared to the CPU velocity, natural dramatics change in the processing technology is embedded significantly, considering the core number [13]. This form of shift is evident in the processor lead that enhances the development of parallel computing. An actual-time application such as SNA, navigation, internet searches, finance and data timeliness necessitates continued computing. The purpose of visualizing information is to avail them fundamentally with the application of several methods centred on the graphic theory. Graph visualization recommends the connection between data and precise interpretation [13]. Nonetheless, online marketing place such as e-bay, Amazon, Flipkart reports millions of users which is data representing a million products being sold in an organization every month.

The process of organizational analysis of data is a good example of how big data works. Up to now, some organizations make use of the tool Tableau for massive visualization of data. This has the capacity to transit complex and large information sets into intuitive pictures. As a result, this enables companies to categorically visualize researches that are relevant to the process of enabling the monitoring of the latest data from clients. This information from customers includes feedback and sentiments about the services being offered to them [14]. Nonetheless, the present massive data visualization tools significantly have projected a lot of issues for the advancement of software and hardware that add to the parallel computing, distributed computing, cloud computing, scalability and visualization procedures. To mitigate this concern, we require a correlation of more mathematics frameworks to PC technology.

### 2.4 Data Security

In the analysis of massive data an incredible amount of information is analysed, mined and correlated to follow a meaningful pattern. All the procedures followed have various policies meant to protect sensitive details. The process of safeguarding sensitive data is a key concern in the massive data analysis process. In this, there is an incredible risk linked to the massive data which makes data security to become a massive data analytics issue. The privacy of the massive data can significantly be developed by the application of methods of authorization, encryption and authentication. Different privacy techniques that massive data application encounters are SNA scales, actual-time privacy monitoring, various device management and lack of data intrusion frameworks [15]. The data privacy issue as a result of massive data has significantly attracted the attention of data security. In that case, attention is designated to formulate the multiple levels of privacy standards and the data prevention framework. Irrespective of the incredible efforts being entitled to safeguard big data, significant improvement is essential. The key issue is to formulate multiple-level privacy and security preserved information systems for massive data.

### 3. Analysis of The Available Research Problems

Massive data analytics and information science are significantly becoming the research focal aspect in

Academia and industries. Information science focusses on analysing massive data and skillset extraction from the information. As such, the application of data science and massive data include data science, modelling, uncertainty, machine learning, uncertain data analysis, pattern recognition, statistical education, signal management and information warehousing. Incredible incorporation of analyses and technologies will stimulate the prediction of the likely drift of events. The vital aspect of this part is to consider the extensive analysis of three significant classes: Internet of Things (IoT), Bio-Inspired Computing, Quantum Computing and Cloud Computing. Nonetheless, this part does not incline to these problems. Instead, it is related to the health database with massive data which can be obtained from analyses [16].

### 3.1 Internet of Things (IoT) for Big Data

International relations have been restructured from the advent of technology, whereas the culture of organizations, revolutions and incredible numbers of individual, organizational cultures are maintained. Presently, machines used in industries are subjected to handling and controlling many automated devices that are driven by technology, hence facilitating the development of IoT. In that regard, the appliances are significantly becoming the major users of the web just like humans with internet browsers. IoT is gaining significantly more attention to new data analytics to comprehend the most promising challenges and opportunities. This avenue of technology provides an important societal and economic implication for future data construction, information technology and the SNA segment. The different regulations for the future projection of the internet will certainly be linked up and controlled smartly by the users. The ideology of IoT is considerably becoming pertinent to the actual world considering the advancement of mobile devices, ubiquitous information methodologies, embedded technologies, data analytics and cloud computing. Apart from that IoT avails the problems in consideration of the velocity, volume and variety.

In essence, just as the internet IoT allows the internet devices to be available in the devices considered in the myriad places and aid the incorporation of device application that range from inconsequential to significant. Converse to that it is inexplicable to comprehend the whole aspect of IoT effectively, including its differences, content and definitions from several similar concepts. Several diversified technologies like computational intelligence and massive data can be interlinked together to enhance the management of information and skillset discovery of massive scale automatic applications. A lot of research aims in this avenue has been highlighted [17]. Skillset acquisition from IoT information set is considered as a significant concern which the massive data experts are constantly facing. In that case, it is fundamental to formulate the data infrastructure to evaluate IoT information. An IoT device produces continuous data streams and the research which are meant to produce tools with precise and sensible data from the datasets with the application of machine learning methodologies. Comprehending these information streams from the IoT devices and evaluating them to obtain sensible data is a problematic concern which leads to massive information analytics. Machine learning algorithms and computing

intelligence method is a key solution meant to deal with massive data from IoT perspective. Major advancements which are linked to the IoT are evaluated in multiple pieces of research done [18].

## 4. Tools for Massive Data Management

Massive tools are present in the process of managing data. In this part, some of the present techniques meant to analyse and manage massive data will be analysed concerning three important tools: Storm, Apache Spark and MapReduce. A lot of these modern tools are centred on the batch process stream management and the interactive data evaluation. These batch management tools are centred on the Apache Hadoop structure like Dryad and Mahout. The streaming of information applications is incredibly utilized for the actual-time data analytics process. Several examples of massive-scale streaming platforms include Splunk and Storm. The interactive evaluation process permits users to incredibly interact in actual time to enhance the process of data evaluation. For instance, Apache Drill and Dremel are massive data platforms which are meant to support an interactive evaluation [19]. These data tools are essential since they enable users to formulate massive data projects. Several researchers [20] also evaluate an incredible list of massive data techniques and tools.

### 4.1 Apache Hadoop and MapReduce

The Apache Hadoop and MapReduce are among the most incredible software applications that have been developed for the analysis of massive data. This software includes Hadoop kernel, Hadoop, MapReduce and Apache hive. MapReduce represents the programming framework used to process significant sets of data that are designated to divide and conquer methodology. The technique is applied following two crucial steps which include map steps and reduce steps. Hadoop operates on two forms of nodes like the worker nodes and master nodes. In the master nodes, the users consider the divisional input into minor sub-issues before distribution to the worker node in the mapping step. In that case, the master nodes link the outputs from the sub-issues in the reduced steps. Apart from that MapReduce and Hadoop operate as an incredible software system used in the process of mitigating massive data issues. It is also essential in checking fault tolerance storage and significant storage of information processes.

### 4.2 Apache Mahout

The Apache Mahout application focusses on the provision of an economical and scalable machine learning methodology for the extensive scale and smart information evaluation application. The key algorithm of this application is centred on pattern mining, classification, clustering, size reduction, regression, batch-based filtering and evolutionary algorithms in addition to the Hadoop platform based on the application of the MapReduce system. The objective of Mahout is to establish a responsive, vibrant and diversified avenue that facilitates the discussion of projects and potential data application instances. The core objective of Apache mahout is to avail necessary tools that will be used to mitigate massive problems. The various industries that have so far applied the scalable machine learning methodology include Amazon, IBM, Yahoo, Facebook and Twitter [21].

### 4.3 Apache Spark

Big data experts consider the Apache Spark as an available source of a large data management system that is established for speed management and sophisticated evaluations. This is incredibly effective to utilize from its introduction in 2009 in the Amp laboratory before being open-sourced in 2010 as the Apache task [22]. This software makes it easier for users to write data application in python, scala and java. Moreover, to map-reduce various operations the software supports the streaming of information, SQL queries, graph data management and machine learning processing. Spark operates from the present Hadoop distributed file scheme structure to assure advanced data functionalities. The driving programs act as the amateur segment of application execution of the spark clusters.

### 5. Projected Research Works

The data amounts that have been retrieved from different application globally and in various fields are purposed to double or triple after two years. The doubling or tripling of data is irrelevant if it is not analysed to obtain credible data. This expert necessity required the advancement of methods that can be utilized to facilitate extensive analysis of data. The advancement of powerful PC systems facilitates the advent of applying these methodologies that stimulates the invention of automatic systems [23]. The transition of information into skillsets should not be considered as a simple task if companies have to witness high-performance massive-scale information processing. Note that this includes the exploitation of parallelism of the present and future PC systems for data mining processes. Apart from that this information might include uncertainties in various forms. A lot of models such as rough sets, neural networks, fuzzy sets and several data generalizations, including hybrid frameworks are retrieved from more than two system models realized due to information representation.

Big data is inclined to incorporate significant features that are essential from a certain viewpoint or dependent on the area of application. In that case, the reduction or inclination methods should be established. Typically, information being collected includes some form of missing valuation which requires generation or tuples characterized by the absent values being eliminated from the information sets before data evaluation. Significantly, the novel problems might include or deteriorate the efficiency, scalability and performance of the established datasets in the intensive computing frameworks. The approaches to be executed significantly leads to the forfeiture of data which renders it as preferred. As such, this project a lot of research problems in an organization and academia in the form of accessing and capturing information incredibly [24]. Rapid processing aims, whereas attaining incredible data performance and high throughput is essential to enhance data storage for future applications. Moreover, algorithmic knowledge for massive data analysis is a fundamental concern that should be mitigated. Addressing the information accessibility necessities of software and the integration of programming languages abstraction is meant to exploit parallelism which is data projection to be investigated [25]. ML tools and concepts are continuously being popular among Academia and researchers to foster

useful results from the various concepts. Research efforts in this part of ML for massive data have concentrated on data management, algorithm optimization and implementation. A lot of ML tools for massive data that have been incorporated recently to necessitate a significant transition and to adopt projected data analytics avenues. While ML tools are advantageous with individual issues, effective tools should be established to mitigate the current massive data problems. The tools should be able to deal with the imbalance and noisy data, inconsistency, wrong valuation and uncertainties.

### 6. Conclusion

Over the past few years data is produced radically. As such processing and evaluating these sets of data are problematic if not proper tools are used. The research challenges, tools and issues have been analysed in this paper to signify how it is essential to establish different measures that will mitigate the inconsistency, uncertainly and noisy data. From this research, it can be concluded that large data systems are defined in their avenues. While some of the datasets are established for batch management some are effective for actual-time information analytic that is to imply that every massive data avenue is designed to serve certain functionality. The various methodologies utilized in the process of data analysis include statistical evaluation, data mining, machine learning, intelligent analysis, quantum computing and cloud computing. In that case, it is expected that future research efforts will focus extensively on the methods to mitigate the evidence of the issue in big data analytics.

### References

- [1] S. Kumar and M. Kirthika, "Big Data Analytics Architecture and Challenges, Issues of Big Data Analytics", *International Journal of Trend in Scientific Research and Development*, vol. -1, no. -6, pp. 669-673, 2017. Available: 10.31142/ijtsrd4673.
- [2] S. Kumar and M. Singh, "Big data analytics for healthcare industry: impact, applications, and tools", *Big Data Mining and Analytics*, vol. 2, no. 1, pp. 48-57, 2019. Available: 10.26599/bdma.2018.9020031.
- [3] F. Ablayev, M. Ablayev, J. Huang, K. Khadiev, N. Salikhova and D. Wu, "On quantum methods for machine learning problems part I: Quantum tools", *Big Data Mining and Analytics*, vol. 3, no. 1, pp. 41-55, 2020. Available: 10.26599/bdma.2019.9020016.
- [4] K. Madhavan and M. Richey, "Problems in Big Data Analytics in Learning", *Journal of Engineering Education*, vol. 105, no. 1, pp. 6-14, 2015. Available: 10.1002/jee.20113.
- [5] "Intelligent analytics for big video data", *Multimedia Tools and Applications*, vol. 79, no. 7-8, pp. 4637-4637, 2020. Available: 10.1007/s11042-019-08540-x.
- [6] J. Bacardit, P. Widera, N. Lazzarini and N. Krasnogor, "Hard Data Analytics Problems Make for Better Data Analysis Algorithms: Bioinformatics as an Example", *Big Data*, vol. 2, no. 3, pp. 164-176, 2014. Available: 10.1089/big.2014.0023.
- [7] A. Singh, "Customized biomedical informatics", *Big Data Analytics*, vol. 3, no. 1, 2018. Available: 10.1186/s41044-018-0030-3.
- [8] A. Hussain and A. Roy, "The emerging era of Big Data Analytics", *Big Data Analytics*, vol. 1, no. 1, 2016. Available: 10.1186/s41044-016-0004-2.
- [9] P. Merritt, H. Bi, B. Davis, C. Windmill and Y. Xue, "Big Earth Data: a comprehensive analysis of visualization analytics issues", *Big Earth Data*, vol. 2, no. 4, pp. 321-350, 2018. Available: 10.1080/20964471.2019.1576260.
- [10] "Healthcare Data Management and Analytics using Big Data Tools", *International Journal of Innovative Technology and*

- Exploring Engineering*, vol. 8, no. 12, pp. 3725-3728, 2019. Available: 10.35940/ijtee.12658.1081219.
- [11]D. P. and K. Ahmed, "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools", *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 2, 2016. Available: 10.14569/ijacsa.2016.070267.
- [12]"A generic data analytics system for manufacturing production", *Big Data Mining and Analytics*, vol. 1, no. 2, pp. 160-171, 2018. Available: 10.26599/bdma.2018.9020016.
- [13]S. Kumar and M. Kirthika, "Big Data Analytics Architecture and Challenges, Issues of Big Data Analytics", *International Journal of Trend in Scientific Research and Development*, vol. -1, no. -6, pp. 669-673, 2017. Available: 10.31142/ijtsrd4673.
- [14]K. Strang, "Problems with research methods in medical device big data analytics", *International Journal of Data Science and Analytics*, vol. 9, no. 2, pp. 229-240, 2019. Available: 10.1007/s41060-019-00176-2.
- [15]N. Rahman, "Data Mining Problems Classification and Techniques", *International Journal of Big Data and Analytics in Healthcare*, vol. 3, no. 1, pp. 38-57, 2018. Available: 10.4018/ijbdah.2018010104.
- [16]J. Park, H. Park and J. Park, "Distributed eigenfaces for massive face image data", *Multimedia Tools and Applications*, vol. 76, no. 24, pp. 25983-26000, 2017. Available: 10.1007/s11042-017-4823-6.
- [17]A. Redway, "Management Tools—Project Planning Procedures", *Industrial Management & Data Systems*, vol. 85, no. 910, pp. 7-11, 1985. Available: 10.1108/eb057412.
- [18]R. Paynter, "Data Mashups as Collection Management Tools", *Collection Management*, vol. 36, no. 1, pp. 68-72, 2010. Available: 10.1080/01462679.2011.531682.
- [19]D. Park and S. Park, "E-Navigation-supporting data management system for variant S-100-based data", *Multimedia Tools and Applications*, vol. 74, no. 16, pp. 6573-6588, 2014. Available: 10.1007/s11042-014-2242-5.
- [20]P. Roohan, "Integration of Data and Management Tools Into the New York State Medicaid Managed Care Encounter Data System", *Journal of Ambulatory Care Management*, vol. 29, no. 4, pp. 291-299, 2006. Available: 10.1097/00004479-200610000-00005.
- [21]M. Rundle, "International Personal Data Protection and Digital Identity Management Tools", *SSRN Electronic Journal*, 2006. Available: 10.2139/ssrn.911607.
- [22]N. Goddard, R. Cannon and F. Howell, "Axiope Tools for Data Management and Data Sharing", *Neuroinformatics*, vol. 1, no. 3, pp. 271-284, 2003. Available: 10.1385/ni:1:3:271.
- [23]"Healthcare Data Management and Analytics using Big Data Tools", *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 3725-3728, 2019. Available: 10.35940/ijtee.12658.1081219.
- [24]G. Steinhart, "Libraries as Distributors of Geospatial Data: Data Management Policies as Tools for Managing Partnerships", *Library Trends*, vol. 55, no. 2, pp. 264-284, 2006. Available: 10.1353/lib.2006.0063.
- [25]R. Bose, "Competitive intelligence process and tools for intelligence analysis", *Industrial Management & Data Systems*, vol. 108, no. 4, pp. 510-528, 2008. Available: 10.1108/02635570810868362.