

LITERATURE REVIEW ON BIG DATA, BIG DATA MODELS AND BIG DATA ANALYTICS

Ms. Monali Milind Chaudhari

Indira College of Commerce and Computer Science, Tathawade.

Email: monali.chaudhari@iccs.ac.in

Ms. Deepali Chaudhari

Indira College of Commerce and Computer Science, Tathawade.

Email: deepali.chaudhari@iccs.ac.in

Abstract

Different sources generating a large amount of data is known as Big Data. Due to the heterogeneous nature of Big Data, it become very difficult to manage and store this Big Data using conventional information systems. In today's era organizations need to manage all kinds of data including traditional structured data, semi-structured and unstructured data. Most of these data models are suitable only for structured data. So, a technology 'NoSQL' came into emergence with features and capabilities that deliver the needs of the Big Data technology. Big Data Analytics is an important method of data collection and is widely used for analyzing structured or unstructured data. In this paper we have done the literature review on Big Data, data models used for Big Data and Big data analytics.

Keywords: Big Data, Big Data Models, Big Data Analytics, NoSQL

Introduction

In the survey it was observed that Social Media, Mobile Devices, Internet Transactions, Networked Devices and Sensors, etc. are the different sources generating a large amount of data known as Big Data. Due to globalization the data sources that generate data for an organization are available in different variety, velocity and volume. Big Data has to deal with two issues- growing size of data sets and increasing of data complexity [43]. Due to the heterogeneous nature of Big Data, it become very difficult to manage and store this Big Data using conventional information systems. Also, the data available from outside is too voluminous and raw for

Conventional information systems to process and use.

By some estimates, data assets greater than 90% is unstructured. Conventional information systems such as relational databases, network and object-oriented databases can process and store structured data. In today's era organizations need to manage all kinds of data including traditional structured data, semi-structured, unstructured and poly-structured data, and the content such as e-mails, web-page content, video, audio, etc.

Traditional database management systems were using different data models to store and manage data. Most of these data models are suitable only for structured data. Also, while designing schema for a structured data it should be known in advance. Hence, when the Big Data came into existence it became a tedious job to scale and manage this Big Data. So, a technology 'NoSQL' came into emergence with features and capabilities that deliver the needs of the Big Data technology. The structured data is disorganized, less described, and in formats poorly suited to long-term reuse [19]. So, there should be proper data model used to store this heterogeneous data.

Big Data Analytics is an important method of data collection and is widely used for analyzing structured or unstructured data. Many organizations were using data mining algorithms for analyzing data stored in data warehouses. When big data and analytics on big data came, the picture changed. Big Data analytics can lead to more effective marketing, new revenue opportunities, improved operational efficiency, competitive advantages over rival organization and other business benefits.

Table 1: Summary of Survey on Big Data

Sr. No.	Title of Paper	Author	Year	Review
1	A survey on data storage and placement methodologies for Cloud-Big Data ecosystem	Yiannis Verginadis, Somnath Muzumdar, Daniel Seybold, Kyriakos Kritikos	2019	Over the time, the type of applications has evolved from batch, compute or memory intensive applications to streaming or even interactive applications. As a outcome, there is a highest rate that applications are becoming complex and has a long duration. Such applications might require frequent-access to multiple distributed data sources. During application deployment and provisioning, the user can face various issues such as (i) where to effectively place both the data and the computation; (ii) how to achieve required objectives while reducing the overall application running cost. [39]
2	A Review of Polyglot Persistence in the Big Data World	Pwint Phyu Khine and Zhaoshun Wang	2019	Big data refers to datasets whose size is beyond of typical database software tools to capture, store, manage, and analyse. Initial big data challenges have been predicted from different viewpoints since the early days of big data. The CAP (Consistency, Availability, Partition Tolerant) conjecture explains why SQL is not sufficient for a distributed system. NoSQL databases are not perfect, in that most of them are evolving technologies. Data coming from various sources is unstructured and it requires analytics to find exact requirement of users. Relational SQL database lives in a shared environment and manipulates data stored in them based on the ACID characteristics. [40]
3	Big Data Processing Technologies in	Nataliya Shakhovska, Nataliya Boyko, Yevgen	2019	NoSQL, Hadoop and Apache Spark are the technologies used

	Distributed Information Systems	Zasoba, Eleonora Benova		for Big Data Processing. Hadoop and Apache Spark are efficient technologies used in Distributed Information Systems[5]
4	Storage solutions for big data systems: A qualitative study and comparison	Khan, Samiya, Xiufeng Liu, Syed Arshad Ali, and Mansaf Alam	2019	This research paper provides feature analysis of 80 NoSQL solutions to facilitate decision making. Moreover, it has been found that many solutions support multiple data models, making the classification categories insufficient. different technologies in a heterogeneous technological environment need to share data to operate. The most optimized method for such data sharing is the use of common data file formats. This research paper categorizes file formats into text-based, row-based, column-based, in-memory and data storage services[44].
5	A Survey on Graph Database Management Techniques for Huge Unstructured Data	Patil N.S.,Kiran P, Kavya N.P., Naresh Patel K.M.	2018	Manages huge unstructured graph databases, The survey of graph data is conducted with the help of data store efficiency, database indexing method, graph indexing method, subgraph matching method. [41]
6	A Big Data smart library recommender system for an educational institution	Aleksandar Simović	2017	A Big Data smart library recommender system is designed using Hadoop Ecosystem which is running slow in parallel processing and there is a need of proper big data model in Hadoop ecosystem is required to process data [13].
7	Big Data Research in Information Systems: Toward an Inclusive Research Agenda	Ahmed Abbasi McIntire	2016	Big data has received considerable attention from the information systems (IS) discipline over the past few years, with several recent commentaries, editorials, and special issue introductions on the topic appearing in leading IS outlets. [38]
8	Exposing Library Data with Big Data Technology: A Review	Chunning Wang, Shaochun Xu, Lichao Chen, Xuhui Chen	2016	This paper proves that library data can be treated as Big Data. Hence big data helps libraries to become more efficient and make innovations in user services.

				Also data analysis should to done properly to identify patterns which can helpful to improve customer satisfaction.[19]
9	An Architecture for Big Data Analytics	Joseph O., Chan	2013	Characteristics of big data, architecture of big data analytics. big data technology moves away from traditional data management approaches. The new paradigm moves towards NoSQL, open source software's and commodity servers. Focusing on running batched and real time analytics on Hadoop platform. [42]

Table 2: Summary of Survey on Data Models used in Various Application Domains

Sr. No.	Title of Paper	Author	Year	Data Models Used	Review
1	Insights from Student Solutions to MongoDB Homework Problems	Ridha Alkhabaz, Seth Poulsen, Mei Chen, Abdussalam Alawini	2021	MongoDB, NoSQL Document Data Model	This paper focuses on how students make errors while executing MongoDB queries given as homework. NoSQL databases here are not used as storage but as learning process [1]
2	A Workload-Driven Framework for NoSQL Data Modeling and Partitioning	Ali Davoudian	2021	No specific	It first generates a generic NoSQL logical schema from the conceptual model and query workload of the system. Then it converts the generic schema to the NoSQL data models regarding their important features and design trade-offs between the schemas of targeted read query performance and storage overhead or consistency maintenance.[2]
3	Secure NoSQL for the Social Networking and E-Commerce Based Big data Applications Deployed in Cloud	Sangeeta Gupta, Narsimha Gugulothu	2018	Cassandra, NoSQL Column-family Data Model	Cassandra, a column-family NoSQL database is used in this paper. Also, a secure framework is designed to provide security to Cassandra in cloud platform. Web crawling is used for extracting data from Social Networking and E-Commerce Based Big Data Applications[10]
4	“(Un)Structuring for the Next Generation: New	Matthew D. Harrington, Dennis B. Christman	2018	Arango DB, NoSQL multimodel	Arango DB, a multimodel NoSQL database is used as a new solution for storing library Data. AQL language is preferred to query data from collections. [11]

	Possibilities for Library Data with NoSQL				
5	A Coherent Healthcare System with RDBMS, NoSQL and GIS Databases	Huanmei Wu, Ashish ambavane, Sunanda Mukherjee, Songan Mao	2017	MongoDB, NoSQL Document Data Model, GIS and MySQL Relational Data Model	A hybrid application for integration of MongoDB, GIS and MySQL databases is proposed for Electronic Health Records (EHRs) and Clinical Decision Support Systems (CDSS) [16]
6	KVFS: An HDFS Library over NoSQL Databases	Emmanouil Pavlidakis, Stelios Mavridis, Giorgos Saloustris, Angelos Bilas	2016	HBase NoSQL Column Family Data Model	NOSQL key-value data store such as HBase is used along with HDFS [20]
7	Enhancing Query Performance of Library Information Systems using NoSQL DBMS: Case Study on Library Information Systems of Universities Indonesia	Hermansyah, Yova Ruldeviyani, Rizal Fathoni Aji	2016	MongoDB, NoSQL Document Data Model, and MySQL Relational Data Model	Query Performance of MongoDB and MySQL databases are compared using Case Study on Library Information Systems of Universities Indonesia. It has been found that MongoDB has better performance than MySQL [21]
8	A NoSQL-Based Framework for Managing Home Services	Marinette Bouet, Michel Schneider	2016	MongoDB, NoSQL Document Data Model	MongoDB NoSQL Document database is used for medical care home services. A application framework is built for handling services like CRUD, computing, display, analysis and archiving [22]
9	A Selection Method of Database System in Big data Environment: A Case Study From Smart Education Service in Korea	Jong Sung Hwang, Sangwon Lee, Yeonwoo Lee, and Sungbum Park	2015	NoSQL Document Data Model, NoSQL Column Family Data Model, NoSQL Graph Data Model	Polyglot persistence used for smart education service. Also, they explained about NoSQL can be considered as efficient database in big data as well as in cloud environment. In different NoSQL data stores explained the key-value data stores are not considered. [24]
10	Applying NoSQL	Reem Qadan Al Fayez, Mike Joy	2015	NoSQL RDF Triple store	NoSQL RDF Triple store is used in this paper. To transform disjoint

	Databases for Integrating Web Educational Stores - An Ontology-Based Approach				heterogeneous web databases entries into a single linked dataset was a challenging task and was achieved by using different ontologies such as MeSH and SNOMED CT [25]
11	NoSQL Databases and Data Modeling Techniques for a Document-oriented NoSQL Database	Robert T. Mason	2015	MongoDB, NoSQL Document Data Model	In this research paper MongoDB NoSQL database is used. They have stated in this paper that Document databases and RDBMS are having same data modeling at conceptual level, but are different at physical level. It is been proved that foreign key concept in RDBMS can be implemented in document databases using embedding. [26]
12	A Digital Library for Plant Information with Performance Comparison between a Relational Database and a NoSQL Database (RDF Triple Store)	Md Arman Ullah	2015	NoSQL RDF Triple store and MySQL Relational Data Model	To store plant information in digital library, MySQL and RDF triple store databases are used. Performance evaluation of both databases is done using plant information. [27]
13	Evaluation of NoSQL databases for EHR systems	Mehmet Zahid Ercan, Michael Lane	2014	No specific	In this paper they have proposed the advantages of using NoSQL databases in Electronic Health Record systems [29]
14	NoSQL and SQL Databases for Mobile Applications. Case Study: MongoDB versus PostgreSQL	Marin Fotache, Dragos Cogean	2013	MongoDB, NoSQL Document Data Model, and PostgreSQL Relational Data Model	MongoDB and PostgreSQL are compared based on major differences between SQL and NoSQL data stores in terms of deployment, data model, schema design, data definition and manipulation. [30]

Table 3: Summary of Survey of Data Analytics used in various Application Domains

Sr. No.	Title of Paper	Author	Year	Data Analytics Algorithms Used	Review
1	Creating a	Nandita S.	2021	GIS,	The data analytics used in library

	Data Science Framework: A Model for Academic Research Libraries	Mani, Michelle Cawley, Amanda Henley, Therese Triumph & Joe M. Williams		visualization, impact measurement, and text and data analysis.	include integrating data analysis into research methodology, creating data visualizations for faculty publications, and partnering with campus units to demonstrate impact through visualizations, including the use of bibliometric analysis. [3]
2	Toward Effective Planning and Management Using Predictive Analytics Based on Rental Book Data of Academic Libraries	Naeem Iqbal , Faisal Jamil , Shabir Ahmad , And Dohyeun Kim	2020	Deep Neural Network (DNN), Support Vector Regressor (SVR), and Random Forest (RF)	The proposed model consists of library data analysis and prediction modules. Data mining techniques are used to analyze and extract useful underlying patterns from library rental book data, which can lead to plan and manage library resources effectively. Secondly, a novel prediction model is proposed based on Deep Neural Network (DNN), Support Vector Regressor (SVR), and Random Forest (RF) to predict future usage of the academic libraries rental books. The performance results of the implemented regression models are evaluated in terms of MAE, MSE, and RMSE. In this paper, it is found that compared to SVR and RF, the DNN model is efficient.[33]
3	A hybrid IT framework for identifying high-quality physicians using big data analytics	Yan Yea, Yang Zhao, Jennifer Shang, Liyi Zhang	2019	similarity match, sentimental analysis, multi-level regression model	A hybrid IT model is proposed to recommend online physicians based on signaling theory by combining similarity matches, reviews, service quality, and basic profiles of physicians using big data. [8]
4	Big Data Analytics in Medicine and Healthcare	Blagoj Ristevski and Ming Chen	2018	Anomaly detection, clustering, classification, association rules	The open-source distributed data processing platform Apache Hadoop MapReduce can be used. These applications should enable applying data mining techniques to these heterogeneous and complex data to reveal hidden patterns and novel knowledge from the data. [12]
5	An analysis of academic librarians' competencies and skills for implementation of Big Data analytics in libraries: A correlational	Khurshid Ahmad, Zheng JianMing and Muhammad Rafi	2018	Introductory knowledge of big data analytics	This survey paper checks the Big Data analytics competencies and skills of Librarians. It was found that librarians are aware of Big Data analytics but should gain deeper insights in analytics. [33]

	study				
6	An integrated big data analytics-enabled transformation model: Application to health care	Yichuan Wanga, LeeAnn Kung, William Yu Chung Wang, Casey G. Cegielski,	2017	Decision Support, Predictive analytics tools	Decision support is one of the crucial big data analytics capabilities due to its ability to create meaningful clinical reports. Big data predictive analytics tools can provide detailed reporting and identify market trends that allow companies to accelerate new business ideas and generate creative thinking. [17]
7	Big Data Analytics for E-Commerce – Its Impact on Value Creation	Avinash B.M., Akarsha B.M.,	2017	No specific	Through use of big data analytics, consumers' behavior online can be analyzed and consumer's interests can be predicted. This gives E-Commerce firms to optimize their marketing plans in real time by integrating all kinds and sources of data. Using big data and related technologies brings robust features to run E-Commerce businesses efficiently, thus creating value for firms. [18]
8	Using data analytics for discovering library resource insights- Case from Singapore Management University	Ning LU, Rui SONG, Dina HENG, Swapna GOTTIPATI, Aaron TAY	2017	pattern matching, rule matching techniques, text mining	This paper put focus on identifying the patterns of how the university student population utilizes e-resources from their library. They found novel solutions to distil information from millions of proxy log records. A data processing workflow and two visualization concepts – horizon chart and word cloud – has been proposed to analyse on numerous resource providers and user groups [34].
9	A Survey of Big Data Analytics in Banking and Health Care today	P. Siva Kumar, B. Murthujavali, Surya PoguJayanna	2016	Data Analytics using Hadoop Map-Reduce and similarity match using HIVE	Hadoop will help in effective citizen care management which can be achieved by providing services to citizens. Data analytics helps to predict the citizen's need by analyzing different classes of citizens. [23]
10	The Study of Big Data Analytics in E-Commerce	Pavithra B 1, Dr.Niranjanmurt hy M2, Kamal Shaker J3, Martien Sylvester Mani F3,	2016	No specific (Survey on machine learning algorithms done)	This research paper discussed the major challenges associated with big data and big data analytics used in e-commerce. Also they discussed solution to the challenges. [35]
11	Analysis of	Mohammad	2015	Advanced	This survey paper focused on

	Research in Healthcare Data Analytics	Ahmad Alkhatib, Amir Talaei-Khoei, Amir Hossein Ghapanchi		Data Visualization (ADV), Presto Hive, Vertica, Key Performance Indicators (KPI) Online Analytics Processing (OLAP), Online Transaction Processing (OLTP), The Hadoop Distributed File System (HDFS) Casandra File System (CFS), Map Reduce Complex Event Processing (CEP), Text Mining Cloud Computing, Mahout, JAQL, AVRO	searching relevant paper in seven popular databases and has listed some data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical operations, reports, decision making, and prediction and prevention system. [28]
12	Library Assessment and Data Analytics in the Big Data Era: Practice and Policies	Hsin-liang Chen, Philip Doty, Carol Mollman, Xi Niu, Jen-chien Yu, Tao Zhang,	2015	keyword search on library logs	Data analytics was used to find the cost of subscription of journals and to specify budget proposals [36]
13	An Introduction To Social Network Data Analytics	Charu C. Aggarwal	2011	Text Mining, Data Mining, Multimedia Information Network Analysis	Social networks contain a tremendous amount of content and linkage data which can be useful for analysis. This linkage data is the graph structure of the social network and the communications between entities; whereas the content data contains the text, images and other multimedia data

					in the network. The richness of this network provides unprecedented opportunities for data analytics in the context of social networks. [31]
14	Web analytics in library practice: Exploration of issues	Anindita Paul,Sanda Erdelez	2010	Web analytics	Web analytics was used improve the quality of services offered to their library users and can add much desired empirical evidence to the process of decision-making. Also web analytics is used to interpret user's behavior on library website. [37]

Major observations from Literature Survey

From the literature survey, it is been identified that a large amount of Big data generated from various sources is unstructured. Big data faces challenges to store this unstructured Data. Also, Relational SQL database manages data based on the ACID characteristics but Big data uses CAP theorem. The tools and techniques used for Big Data was observed and we identified that Hadoop ecosystem and NoSQL are frequently used in Big Data. We have also observed that Big data research has been done more in domains of Healthcare, agriculture, e-commerce but there is scope of research in other domains such as library, education, etc.

The survey of Data models used in various Big Data Application Domains suggested that there is no one data model which can be suitable for all Big Data Applications. From the literature survey, it was found that MongoDB, NoSQL Document Data Model is gaining most popularity. Because of all heterogeneous data format, it was necessary to store data in different database management systems and then combine all in one application. But this may need separate installation of each DBMS server, and further join this data at distinct places. Hence, it is more advantageous to build a single multi-model database. This can also be achieved by polyglot persistence. Having a single database platform for relational and NoSQL data is beneficial to users; this approach reduces significantly integration, migration, development, maintenance, and operational issues.

Data analytics survey shows which machine learning tools and techniques can be used to extract meaningful information from various domains. Big data predictive analytics tools can provide detailed reporting and identify market trends that allow companies to accelerate new business ideas and generate creative thinking[17] Also, Big data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical operations, reports, decision making, and prediction and prevention system. [28] Moreover, Data analytics helps to predict the user's need by analyzing different classes based on Classification Machine Learning Algorithm [23]. Text search or keyword search is one of Data analytics technique used in social media related research papers.

Research Conclusion

It is observed that most of the work for Big Data and Data Analytics is done in healthcare, banking, e-commerce, etc. can use this Big Data to enhance decision-making skills and in user behavior. Also, from study of data models used for Big Data Applications shows that there is a need of multi-data model. Hence, we need to propose a multi-data model suitable for the data for different organization. Data Analytics survey proves that using various machine learning algorithms and techniques can help organization to find new revenue opportunities, and expand its user and services.

References

- [1] Alkhabaz, R., Poulsen, S., Chen, M., & Alawini, A. (2021, June). Insights from student solutions to mongodb homework problems. In *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1* (pp. 276-282).
- [2] Davoudian, A. (2021). *A Workload-Driven Framework for NoSQL Data Modeling and Partitioning* (Doctoral dissertation, Carleton University).
- [3] Mani, N. S., Cawley, M., Henley, A., Triumph, T., & Williams, J. M. (2021). Creating a data science framework: A model for academic research libraries. *Journal of Library Administration*, 61(3), 281-300.
- [4] Anna, N. E. V., & Mannan, E. F. (2020). Big data adoption in academic libraries: a literature review. *Library Hi Tech News*.
- [5] Shakhovska, N., Boyko, N., Zasoba, Y., & Benova, E. (2019). Big data processing technologies in distributed information systems. *Procedia Computer Science*, 160, 561-566.
- [6] Al-Barashdi, H., & Al-Karousi, R. (2019). Big Data in academic libraries: literature review and future research directions. *Journal of Information Studies and Technology*, 2018(2), 13.
- [7] Ball, R. (2019). Big data and their impact on libraries. *American Journal of Information Science and Technology*, 3(1), 1-9.
- [8] Ye, Y., Zhao, Y., Shang, J., & Zhang, L. (2019). A hybrid IT framework for identifying high-quality physicians using big data analytics. *International Journal of Information Management*, 47, 65-75.
- [9] Kaladhar, A., Naick, B. D., & Rao, K. S. (2018). Application of big data technology to library data: a review. *International Journal of Library and Information Studies*, 8(2), 25-30.
- [10] Gupta, S., & Gugulothu, N. (2018). Secure NoSQL for the social networking and e-commerce based bigdata applications deployed in cloud. *International Journal of Cloud Applications and Computing (IJCAC)*, 8(2), 113-129.
- [11] Harrington, M. D., & Christman, D. B. (2019). (Un) Structuring for the Next Generation: New Possibilities for Library Data with NoSQL.
- [12] Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. *Journal of integrative bioinformatics*, 15(3).
- [13] Simović, A. (2018). A Big Data smart library recommender system for an educational institution. *Library Hi Tech*, 36(3), 498-523.
- [14] Li, J., Lu, M., Dou, G., & Wang, S. (2017). Big data application framework and its feasibility analysis in library. *Information Discovery and Delivery*, 45(4), 161-168.
- [15] Golub, K., & Hansson, J. (2017). (Big) data in library and information science: a brief overview of some important problem areas. *Journal of universal computer science (Online)*, 23(11), 1098-1108.
- [16] Wu, H., Ambavane, A., Mukherjee, S., & Mao, S. (2017). A coherent healthcare system with RDBMS, NoSQL and GIS databases.
- [17] Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, 55(1), 64-79.
- [18] Avinash, B. M., & Akarsha, B. M. (2017). Big data analytics for e-commerce—its impact on value creation. *International Journal of Advanced Research in Computer and Communication Engineering*, 6(12), 181-188.
- [19] Wang, C., Xu, S., Chen, L., & Chen, X. (2016, June). Exposing library data with big data technology: A review. In *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)* (pp. 1-6). IEEE.
- [20] Pavlidakis, E., Mavridis, S., Saloustros, G., & Bilas, A. (2016, April). KVFS: An HDFS Library over NoSQL Databases. In *CLOSER (1)* (pp. 360-367).
- [21] Ruldeviyani, Y., & Aji, R. F. (2016, October). Enhancing query performance of library information systems using NoSQL DBMS: Case study on library information systems of Universitas Indonesia. In *2016 International Workshop on Big Data and Information Security (IWBIS)* (pp. 41-46). IEEE.

- [22] Bouet, M., & Schneider, M. (2016). A NoSQL-based framework for managing home services. *Open Journal of Information Systems (OJIS)*, 3(1), 1-28.
- [23] P. Siva Kumar, B. Murthujavali, & Surya PoguJayanna, (2016). A Survey of Big Data Analytics in Banking and Health Care today. *International Journal of Computer Engineering In Research Trends*, Volume 3, Issue 12.
- [24] Hwang, J. S., Lee, S., Lee, Y., & Park, S. (2015). A selection method of database system in Bigdata environment: a case study from smart education service in Korea. *International Journal Advance Soft Computing Application*, 7(1), 9-21.
- [25] Al Fayez, R. Q., & Joy, M. (2015). Applying NoSQL Databases for Integrating Web Educational Stores-An Ontology-Based Approach. In *Data Science: 30th British International Conference on Databases, BICOD 2015, Edinburgh, UK, July 6-8, 2015, Proceedings 30* (pp. 29-40). Springer International Publishing.
- [26] Mason, R. T. (2015, July). NoSQL databases and data modeling techniques for a document-oriented NoSQL database. In *Proceedings of Informing Science & IT Education Conference (InSITE)* (Vol. 3, No. 4, pp. 259-268).
- [27] Ullah, M. A. (2015). A Digital Library for Plant Information with Performance Comparison between a Relational Database and a NoSQL Database (RDF Triple Store).
- [28] Alkhatib, M., Talaei-Khoei, A., & Ghapanchi, A. (2016). Analysis of research in healthcare data analytics. *arXiv preprint arXiv:1606.01354*.
- [29] Ercan, Mehmet & Lane, Michael. (2014). An Evaluation of NoSQL Databases for Electronic Health Record Systems.
- [30] Fotache, M., & Cogean, D. (2013). NoSQL and SQL Databases for Mobile Applications. Case Study: MongoDB versus PostgreSQL. *Informatica Economica*, 17(2).
- [31] Aggarwal, C. C. (2011). *An introduction to social network data analytics* (pp. 1-15). Springer US.
- [32] Iqbal, N., Jamil, F., Ahmad, S., & Kim, D. (2020). Toward effective planning and management using predictive analytics based on rental book data of academic libraries. *Ieee Access*, 8, 81978-81996.
- [33] Ahmad, K., JianMing, Z., & Rafi, M. (2019). An analysis of academic librarians competencies and skills for implementation of big data analytics in libraries: a correlational study. *Data Technologies and Applications*.
- [34] LU, N., SONG, R., HENG, D. L. G., GOTTIPATI, S., & Tay, A. (2017). Using data analytics for discovering library resource insights: Case from Singapore Management University.
- [35] Pavithra B, Dr.Niranjanmurthy M, Kamal Shaker J, Martien Sylvester Mani F, "The Study of Big Data Analytics in E-Commerce", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 5, Special Issue 2, October 2016 Copyright to IJARCCCE DOI 10.17148/IJARCCCE 126
- [36] Chen, H. L., Doty, P., Mollman, C., Niu, X., Yu, J. C., & Zhang, T. (2015). Library assessment and data analytics in the big data era: Practice and policies. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.
- [37] Paul, A., & Erdelez, S. (2009). Web analytics in library practice: Exploration of issues. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1-6.
- [38] Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the association for information systems*, 17(2), 3.
- [39] Mazumdar, S., Seybold, D., Kritikos, K., & Verginadis, Y. (2019). A survey on data storage and placement methodologies for cloud-big data ecosystem. *Journal of Big Data*, 6(1), 1-37.
- [40] Khine, P. P., & Wang, Z. (2019). A review of polyglot persistence in the big data world. *Information*, 10(4), 141.
- [41] Patil, N. S., Kiran, P., Kiran, N. P., & KM, N. P. (2018). A survey on graph database management techniques for huge unstructured data. *International Journal of Electrical and Computer Engineering*, 8(2), 1140.

- [42] Chan, J. O. (2013). An architecture for big data analytics. *Communications of the IIMA*, 13(2), 1.
- [43] Huda, M Misbachul & Hayun, Dian & Martun, Zhin. (2015). Data Modeling for Big Data. *Jurnal ULTIMA InfoSys*. 6. 1-11. 10.31937/si.v6i1.273.
- [44] Khan, S., Liu, X., Ali, S. A., & Alam, M. (2019). Storage solutions for big data systems: A qualitative study and comparison. *arXiv preprint arXiv:1904.11498*.