ISSN: 0067-2904

Exploring the Literature of Data Analytics Services on Cloud Computing: A Comprehensive Summary

Safa S. Abdul-Jabbar[1]*, Alaa k. Farhan[2]
1Computer Science Department/ College of Science for Women University of Baghdad, Baghdad, Iraq
2 Computer Science Department, University of Technology, Baghdad, Iraq

Received: 21/12/2022 Accepted: 15/3/2023 Published: 29/2/2024

Abstract

Data generated from the internet and modern applications is extensive and rapidly expanding. So, all modern applications must successfully perform tasks using this massive data. Therefore, one of the significant factors for the success of any application is understanding and extracting meaningful information using digital analytics tools to positively impact the application's performance and deal with challenges that can be encountered. On the other hand, cloud computing is simply an environment comprising a collection of high-performance services from various vendors. These services can frequently access and process massive amounts of data faster than a traditional computer. One of these services is cloud analytics, which applies cloud-based analytic algorithms against data in a private or public cloud to get the desired outcome. This paper differentiates between several analytics methodologies to create highly consistent, logical, and information-rich summaries. This research has discussed and analyzed many studies related to this field. Therefore, the results of this paper can be utilized to determine the advantages of these methods, which will help future researchers in their research for a more organized and thorough analysis when dealing with such applications. In addition, this paper attempts to discover new directions and propose new guidelines in this evolving field.

Keywords: Data Analytics, Cloud Computing, Analytics Service, Big data, Big data analysis, Big data Literature**.**

استكشاف أدبيات خدمة تحليل البيانات في الحوسبة السحابية: ملخص شامل

صفا سامي عبد الجبار[1]*, علاء كاظم فرحان[2]
[1]قسم علوم الحاسوب، كلية العلوم للبنات، جامعة بغداد، بغداد، العراق
[2]قسم علوم الحاسوب، الجامعة التكنولوجية، بغداد، العراق

الخلاصة

تتجمع البيانات الناتجة عن الإنترنت والتطبيقات الحديثة وتتوسع بسرعة كبيرة. لذلك، ومن المهم ايضاً أن تؤدي التطبيقات الحديثة جميع مهماها بنجاح باستعمال هذه البيانات الضخمة. لذلك، فإن أحد العوامل المهمة لنجاح أي تطبيق هو فهم واستخراج المعلومات ذات المغزى باستعمال أدوات التحليل للتأثير بشكل إيجابي على أداء التطبيق. من ناحية أخرى، الحوسبة السحابية هي ببساطة بيئة تضم مجموعة من الخدمات عالية الأداء من عدة مجهزين مختلفين. ويمكن لهذه الخدمات الوصول بشكل متكرر إلى البيانات الضخمة ومعالجتها بشكل أسرع من الحاسوب التقليدي. إحدى هذه الخدمات هي التحليلات السحابية التي تحدد الخوارزميات التحليلية القائمة

*Email: safa.s@csw.uobaghdad.edu.iq

على السحابة مقابل البيانات الموجودة في السحابة الخاصة أو العامة للحصول على النتيجة المرجوة. تميز هذه الورقة بين العديد من منهجيات التحليلات لإنشاء ملخصات منطقية وغنية بالمعلومات. كما ويناقش هذا البحث العديد من الدراسات المتعلقة بهذا المجال. لذلك، يمكن الاستفادة من نتائج هذه الورقة لتحديد مزايا هذه الأساليب لمساعدة الباحثين المستقبليين في أبحاثهم من أجل تحليل أكثر تنظيماً وشمولية عند التعامل مع مثل هذه التطبيقات. بالإضافة إلى ذلك، تحاول هذه الورقة أيضًا اكتشاف اتجاهات جديدة واقتراح إرشادات جديدة في هذا المجال المتطور.

1. Introduction

The world's data has increased at an exponential rate. A new generation of technologies and architecture that facilitate fast data capture, transport, storage, and analysis for enormous amounts of data enables the data analysis approach. These numerous data sets are now known as “big data” [1]. This term is still evolving to describe any significant volume of structured, semi-structured, or unstructured data that can be used to extract only relevant information. Big data cannot be processed using ordinary databases because the amounts of information are too large for one machine to handle. Big data analytics is a growing area with vast volumes of data to find hidden correlations, relationships, and other findings. The latest advances in computer technologies, techniques, and approaches created to handle massive data have made big data technology conceivable [2]. With the explosion of structured and unstructured data in many research areas such as genetics, earth science, and others, processing, managing, and analyzing trends utilizing standard databases have become impossible because of the variety of data types collected and the effort required for processing. Therefore, in 2013, Yadav and his team published a paper summarizing the design and methods of big data [3]. These methods describe numerous big data structures with several techniques and list several tools used to evaluate the proposed system's performance. Also, this research presents security challenges that can be faced when dealing with a large data set [3].

Due to the diverse applications of Internet of Things (IoT) devices, data collection through networks or the Internet has become a significant part of our everyday routine [4]. An example of the investment in data analytics tools in the cloud environment is to integrate the Internet of Things with big data analytics, which allows real-time tracking of the collected data and makes decisions to ensure consistency and throughput. Furthermore, in the context of healthcare systems, it provides clinicians with real-time observation of the patient's condition, allowing them to choose the best possible treatment option. In addition, this integration will allow for a reduction in therapy latencies and healthcare costs [5]. This article delves into the transformative impact that data analysis services have on cloud computing, especially when handling vast amounts of big data. It will examine how these services can offer numerous advantages and pose certain challenges for businesses operating in various sectors. Also, big data explains how data analytics can help improve the critical functions of big data in the cloud environment (i.e., any cloud application dealing with massive data). Finally, this article provides an overview of cloud computing and big data analytics principles, challenges, and solutions.

2. Literature Reviews

It is important to remember that when discussing big data, the word refers to the trend rather than a specific technology. As a result, more authors define big data by assigning a set of V's rather than attempting to define this issue. These V's are volume, velocity, variety, variability, veracity, visualization, value…, etc. [6-9]. The big data features are one of the primary reasons for the popularity of cloud computing and have led various academic and business initiatives to investigate its capabilities and advancements. Many tools have been provided to handle the

massive data provided in cloud computing environments, such as data analytics. Therefore, the following sections present some examples of how big data analytics are used in various industries:

2.1 Healthcare

Big data is a common topic in almost all areas of research. Big data has many characteristics in healthcare, such as heterogeneity, incomplete information, responsiveness, duration, security, and copyright. These characteristics present several data storage, processing, and transmission challenges for further health-related research [10]. Like other organizations, healthcare organizations aim to organize, process, and analyze data from multiple sources. Big data in the medical domain is derived from hospital information resources, doctors' work, patient care activities, medical tests, medical imaging, magnetic resonance imaging, computer tomography, health histories, pharmacies, and others [11, 12]. Many research papers published dealt with the healthcare field depending on data analytics techniques. For example, in 2018, Jindal et al. published a research paper focusing on cloud computing by providing Healthcare-as-a-Service based on the preliminary cluster construction, extraction, and analysis of vast amounts of data. A fuzzy rule-based classifier is constructed in the proposed system for efficient data classification decision-making. Membership functions are created for fuzzification and defuzzification operations to generate inferences from acquired data. The suggested technique is assessed using several measures. Findings were collected to confirm the effectiveness of the suggested system in a cloud computing environment in terms of several performance evaluation measures [13]. While in 2019, Tamam and his team sought to design and develop a graphical user interface-based application system that can represent the form of heart-rate waves, blood oxygen levels, and temperature. The measurement data generates a result that is saved in the database. Based on the test results, the system can show the heart-rate pattern, pulse oximetry in blood, temperature, and diagnosis [14].

2.2 Telecommunication and Social Media

The emerging sectors of the social web, user and industrial analytics, and governmental monitoring are where privacy-sensitive big data applications can be found [15–17]. For example, mobile service providers encounter various problems, including low mobile service demand and attrition management. Therefore, mobile service providers analyze various factors, including demographic data, customer preferences, household structure, and usage details, to model customer preferences and provide them with relevant, customized services. This is known as “targeted marketing,” which increases mobile service providers' revenue by promoting mobile service adoption and reducing defection [18]. Rehman and Al-Raqom designed applications in 2019 with three organizations that deployed big data systems for various tasks. These organizations require qualified personnel for data management, security, and other applications. The outcomes of this research are encouraging for businesses looking to implement innovative systems and technologies [19].

2.3 Insurance

Like other technologies, big data will continue to be a business enabler. It will provide newer insights (using social media, telematics, etc.) such as tracking and other emerging technologies to enterprises at a rapid speed where big data can be beneficial. To comprehend customers' needs and concerns, insurers develop new products and improve existing ones to ensure efficient and competitive service levels [20]. In 2020, Zheng and Guo designed a new application of big data technology in insurance innovation by defining the insurance challenges, efficiency improvement methods, and a good strategy plan for insurance companies [21].

2.4 Financial Firms and Banking

Banks mostly use big data to improve the transparency and audibility of their organizations [22]. Therefore, many papers discussed big data applications and IoT in banking organizations. For example, Chang et al. explained the potential bank's operations that can generate massive data, then listed the primary usage of big data in the banking organization. Finally, several ethical interests and proper solutions were presented for banking systems [23]. At the same time, Lawler and Joseph produced a scientific paper that covers big data analytics methodology in the financial industry and assists educators in developing big data analytics curricula. In addition, to keep current with the trends of organizations that have successfully used big data analytics systems [24]. On the other hand, many papers were also published to design new applications, such as Ngai et al., who created a classification approach to study advanced analytics in fraudulent financial detection. In this research, they discovered a new text analysis system that used Naive Bayes text analysis to identify people who were likely to commit fraud [25]. Also, Delgosha and his team elucidated data analytics in banking. This research increases the grasp of the critical management difficulties of big data in a dynamic business context by providing further practical steps for both researchers and decision-makers by designing and implementing a four-round Delphi study [26].

2.5 Education

Every day, people use cloud computing in certain forms; the education sector is one that started using the cloud environment to improve its work and make it more flexible. Because teaching methods are evolving and schools are becoming more technology-dependent, it is critical to consider how to incorporate the most recent technologies into teaching and learning methods. As a result of sharing with cloud-based information systems, academic institutions can focus more on providing critical tools to students, teachers, administrators, and employees [27]. Beck and Mostow describe learning decomposition as a method for determining the comparative efficacy of different educational activities. This method is a generalization of learning curve analysis that employs nonlinear regression to determine how to weigh various practice opportunities against one another. The research findings show that one-size-fits-all education is 95% accurate and that the Intelligent Transportation Systems (ITS) community is well-positioned to research what form of instruction is best for each user [28]. Waga et al. propose a government-owned and maintained cloud computing framework. Any solar-powered computer might access the collection and can use primary connections. It is explored how cloud provider decisions are influenced by digital content development capacity. The impact of achieving goals including universal primary education, environmental protection, and illness prevention through research on Kenya's Vision 2030 is examined. Instructor shortages, classroom space constraints, and geographical disparities are all resolved [29].

2.6 Business

In 2011, Agrawal et al. presented baseline research covering both types of systems: those that enable update-intensive applications and those that support ad-hoc analytics and decision support. Then they suggest multiple Map-Reduce frameworks over Hadoop Distributed File System solutions for dealing with big data analysis challenges. Finally, this paper uses Hadoop to build MapReduce techniques for demand analysis and access patterns [30].

The three-aspect method presented by Paweoszek and Wieczorkowski provides a framework for understanding the different components of big data, which can be useful in assessing its impact and potential applications. The three key components are:

- Technological: This aspect of big data refers to the various tools, methods, and technologies used to capture, store, process, and analyze large volumes of data. It includes hardware and software infrastructure, data storage systems, analytics tools, and other related technologies.

- Business: This aspect of big data focuses on its commercial applications, mainly its role in decision support. Businesses use big data to gain insights into customer behavior, market trends, and other relevant factors to make better-informed decisions.
- Social: This aspect of big data relates to the societal impact and ethical considerations of data processing outcomes. It includes issues such as data privacy, security, and the potential for bias or discrimination in data analysis.

When using big data analytics tools, all three criteria should be considered to ensure that data is being used in a responsible and effective manner. This approach can help organizations achieve their goals while also addressing ethical and social concerns related to the use of big data. [31]. Jelonek discusses his thoughts on the topic of big data. This research also shows that big data analytics may be used to help manage a corporation. It also identifies the areas and activities wherein big data analytics can provide the most value to businesses [32]. In 2020, Ali and Farhan utilized encrypted lossless compression technology to develop a novel method for improving the data storage of a quick response code. This algorithm has the potential to simplify marketing and selling processes [33].
There are also some drawbacks in each research sector. Some examples of these drawbacks are shown in Table 1.

Table 1: The most important challenges in each sector in a cloud computing environment

Reference	Research Sector	The challenge
Lee et al. [5]	Healthcare	Many healthcare applications must obtain, maintain, and analyze data in real time. These applications only deal with linear relationships between features and only use baseline features.
Rehman and Al-Raqom[19]	Telecommunication and Social Media	In addition to security and privacy concerns, telecommunications corporations and social media platforms must manage and utilize large amounts of data.
Zheng and Guo [21]	Insurance	Insurance Companies Face Many Challenges, such as (Conflicts, Data Value, Insufficient Interactions, Data Isolation, and Cross-border Competition).
Delgosha [26]	Financial Firms and Banking	The most critical challenges for banking systems are (High prices for acquiring essential technologies, data security, and data integrity).
Ercan [34]	Education	Where some services required to deliver most of the information technology facilities required by users did not exist until now, there are still challenges and restrictions with program offers, service contracts, and increasingly severe security threats. Finally, not all cloud service providers have the same capabilities on technical levels.
Xue and Xin[35]	Business	When dealing with business, there are many challenges, such as issues for decision-makers to consider; cloud network traffic will be detrimental to cloud performance, and there is a risk of loss of valuable data such as clients' data and businesses' sales reports.

3. Big Data Analytics Problems and Issues

Cloud computing has been used in many applications and has many appealing features that make it attractive for future IT software and systems. The National Institute of Standards and Technology has identified five key features of the cloud computing environment: Measurable services, high flexibility, resource pooling, self-service on-demand, and mobility [2] [36]. In addition, many features are provided when using the cloud environment, such as virtualization, multitenancy, configuration, stability, optimized resource consumption, scale economies, and scalability [37, 38]. These features can be applied to the analysis systems to simplify and improve process efficiency and accuracy. On the other hand, the most difficult question is: What if the size of stored data becomes vast, diverse, and no longer manageable? How much data can be handled? Is it necessary to save all data? Is it necessary to analyze all the data? How do you figure out which data points are the most crucial? What are the most straightforward and efficient ways to use most data? [39]. In addition to these critical system design questions, there are some additional challenges that data analytics systems must address when working with big data, such as [10]:

3.1 Privacy and Security

Security is one of the most significant concerns people face throughout their lives. Internal and external attacks can compromise data security [40, 41]. Privacy has been extensively researched in recent decades. Primarily, previous research has focused on encryption, networking, and security. Unfortunately, because of the enormous volume of big data, traditional encryption solutions are challenging to use successfully in this era. Limited processing and storage are other constraints. In addition, the ability of mobile devices for encryption and decryption operations is impossible under big data requirements [42, 43].

3.2 Scalability and Storage Issues

The amount of data available grows dramatically. Data management systems with horizontal scalability have increasingly become popular to meet the challenges of gathering massive amounts of data. Due to the issues in big data management—volume, velocity, variability, validity, and value—data have become a vital commodity in enterprises, and data analysts battle to make optimal business decisions. Due to previously exceptional but rapidly expanding needs such as scalability, data intensiveness, high availability, fault tolerance, and the ability to handle varied data structures, large-scale data analytics has become a high-computational concept. Parallel computing resources are necessary to meet computational and memory needs when these challenges grow in magnitude [44].

3.3 Processing Power

The scale of big data is the first thing that comes to mind when dealing with big data. Managing enormous and fast-expanding volumes of data has been difficult for decades. In the past, this problem was solved by faster processors, which followed Moore's law and provided us with the needed resources to deal with growing data volumes. However, there is a significant change; for example, data volume grows faster than computational resources, while CPU speeds remain constant [45].

3.4 Timeliness and Heterogeneous Data

One of the essential characteristics of big data is heterogeneity, which causes challenges with data integration and big data analytics [46]. Heterogeneity and responsiveness can be tolerated when there are many data points. The richness and diversity of natural language offer unparalleled depth and nuance, enabling us to express complex ideas and emotions with greater clarity and precision. Whether using metaphor, analogy, or other rhetorical devices, language

provides a rich palette of tools for communication and expression, allowing us to convey abstract concepts and subtle shades of meaning that might otherwise be difficult to express. The sheer variety of vocabulary and grammatical structures in natural language also enables us to adapt our communication to different contexts, audiences, and purposes, enhancing our ability to connect with others and convey our intended message with greater impact. On the other hand, machine analysis techniques, which are designed for homogeneous data, cannot comprehend nuance. As a result, data must be carefully formatted before analysis. Even after data cleaning and mistake correction, there will likely be some incompleteness and errors in the data. During data analysis, these mismatches and mistakes must be controlled. It takes much work to perform the data analysis process correctly [47].

3.5 Skill Requirement

Because big data is still a new and evolving technology, it attracts companies and college students with various new skills. These abilities should not be confined to technical abilities but must include research, analytic, interpretative, and creative abilities. Individuals must gain these talents, which necessitates holding training programs at institutions [39].

On the other hand, there are many technical challenges when working with data analytics. Some of them are listed below: [39] [19]:

1- Out-of-context data

Data reduction is one of the most prevalent methods for fitting into a statistical model. It is crucial to keep context after data extraction because it will lose meaning and significance when taken out of context [48].

2- Fault Tolerance

The critical attribute of big data analytics systems is fault tolerance, which ensures availability, stability, and reliable quality even when there are failures. However, finding an efficient fault tolerance strategy in a big data system is difficult because fault tolerance must conform to specific efficiency and resource limits [46, 47].

3- Continuous Availability

When you rely on extensive data to feed your income and support 24/7 selection processes, high reliability alone is not enough. It's essential to ensure that your data is always recoverable in case of a failure or data loss. This requires robust backup and disaster recovery strategies, as well as effective data management practices. By prioritizing data recoverability, you can minimize the risk of disruptions to your operations and ensure that your business can continue to function smoothly even in the face of unexpected challenges. The system's ability to simultaneously process streaming data and respond to real-time inquiries from thousands of users is restricted. People expect the systems they interact with to respond in real-time or near real-time [19].

4- Quality of Data

Finding the correct data in a flood of data is a difficult task because many sources provide data, including social media websites, journals, etc. Companies need help to find the correct data and determine how to use it effectively. As a result, it is necessary to discover the guidelines that will assist in identifying the correct data [50].

5- Fault Detection

Data complexity, such as high dimensionality, fast-flowing data streams, and strong nonlinearity, makes fault detection applications challenging. From the standpoint of data modeling, excessive dimensionality may result in the well-known “curse of dimensionality,” causing fault detection methods to lose their accuracy. Fast-moving data streams necessitate algorithms that respond in real-time or near real-time when new samples arrive. Because of the

high nonlinearity, fault detection methods must have the appropriate expressive capacity to prevent overfitting or underfitting issues. The majority of known defect detection methods operate in low-dimensional spaces. Therefore, only a few research papers have handled the issues of high dimensionality and data streams. The smoothness, efficacy, robustness, and interpretability of existing nonlinear defect detection techniques must be improved [51].

6- Dealing with Outliers

Visualization, which allows graphical data representations, may explain trends and outliers faster than tables with numbers and text. Outliers often make up 1–5% of the data when working with big data. However, examining 1 to 5% of data when working with enormous amounts is challenging. The most common option is eliminating outliers from the data or constructing a separate chart for them. Outliers may not indicate the data, but they can provide previously unknown and possibly significant information [50].

7- Digital Divide

One of the most significant difficulties is gaining access to massive data. Large amounts of social data are available to data and social media firms. However, only some companies decide who has access to data and how much access they have. Few sell access rights for a high price, while others give researchers access to a subset of data sets. In big data, this leads to a digital divide. In the world of big data, there are three categories of people and organizations. First, people who create data include anyone using the web, mobile devices, or other technology. Second, the group of people who collect data is restricted, while people who can analyze data are in the third group. This is the most minor and protected group [48, 52].

To deal with the above problems, many big data analytic methods were proposed to fulfill their organization's goals. Everyone wants to concentrate on obtaining key values and information from a large dataset. Therefore, the most often utilized big data analysis approaches are [53]:

- Bloom Filter is a collection of hash functions that can be used to filter data. The main idea behind this method is that data hash values are stored in bit arrays.
- Hashing: Data is mutated into the index and numeric values using the hashing process.
- Index: An index is a helpful tool for lowering disk reading and writing costs while enhancing query insertion, deletion, and modification performance. The disadvantage of this strategy is the additional cost of index file storage.
- Trial: is a variant of Hash Tree known as “Tire Tree.”
- Parallel Computing: Parallel computing (in contrast to serial computing) uses several capabilities simultaneously to execute an operation.

4. Discussion and Findings

Big data has become a general term that denotes vast volumes of complex datasets rapidly generated from various sources. This big data potentially has an immense hidden value that needs to be discovered by intelligently utilizing analytics services. Since the emergence of data science and advanced techniques for data manipulation, such as data mining and data analytics, there has been a significant increase in data manipulation. The demand for literature analyses that may provide a complete summary of the benefits and drawbacks of data analytics usage for professionals and academics has grown tremendously.

This paper provides an overview of big data and the processes involved in big data analytics, the importance of cloud computing when employing data analytics, and its current applications. It also discusses various data processing tools and techniques. In addition, it presents many sectors where big data may help improve services. Finally, technological growth and limitations

for future research in improving big data analytics have been highlighted. However, this paper discussed many challenges and issues that must be addressed before accepting and adopting this new technology.
Nevertheless, these challenges and issues will be found in various businesses that depend on this technology to increase the findings of their work. Therefore, various flaws should be analyzed to provide a particular advantage to the cloud computing platform. As presented in Table 2, some crucial issues were discussed that might be raised to allow businesses to consider these issues from the start and devise solutions.

Table 2: The most important issues when using the cloud computing

Reference	The challenge	Solutions'
Hossain [54]	Scalability	the ability to adjust the required resources to evolve user needs through intelligent resource management and effective monitoring.
Sen [55]	Data protection, confidentiality, and confidence	As a result, the data should be encoded and managed securely and efficiently. Security audits and certifications must also be undertaken to strengthen users' security.
Tafani et al. [56]	Energy efficiency	It is also critical to lessen the electric charge by utilizing microprocessors that consume less energy and are adaptive in their application.
Xue and Xin [57]	Provision, flexibility, and recovery	Using redundant systems to ensure continuous service to all users on all days and at all times while avoiding net traffic overload
Berisha [58]	Big data volumes and quality	The rapidly increasing size and heterogeneity of data are the essential issues that should be preserved and dealt with in cloud computing.

5. Conclusion and Recommendations

In this research, we distinguish between several data analytic approaches according to their challenges in cloud computing environments that aid in the construction of internally coherent, simple summaries. As a result of the above mentioned in this paper, we can conclude that cloud analytics is still plagued by numerous issues, such as:

- Loss of central control over the data.
- It requires more experience to deal with each sector and manage data in the most helpful way.
- It required diverse architectures for each sector to meet the organization's requirements.
- Security and Privacy issues.
- Poor data quality makes the service less reliable in making decisions.
- Sequential data processing means delays in obtaining results, despite using cloud environments

This paper will also provide several solutions as suggestions to deal with the most critical challenges mentioned previously in this paper, as shown in Table 3.

Table 3: The most critical challenges and the suggested solutions

Challenge	Suggested Solution
Data Isolation and data Conflicts	Using the Blockchain's technology and hybrid Public Key Infrastructure depending on data analytics to remove redundant data and select only the most recent data.
Security Threats (data confidentiality and data integrity)	Using a hybrid system for encrypting data, such as honeypot encryption hybrid with hash algorithms
Cloud network traffic	Analyze traffic patterns based on machine learning and data mining algorithms
Dealing with different levels of security	Using Blockchain technology to provide processing authentication and data integrity based on associated hashing values with the Twofish encryption algorithm.
Sequential Data Processing	Using the Blockchain behavior in data analytics systems
Big data Issues	Using AI and data analytics techniques to preprocess data, remove redundant data, and select only valuable data from these data.

Also, future cloud analytics research has virtually endless possibilities for all sectors. Therefore, its application is tied to utilizing this service for creating unique conditions based on the intended purpose and outcomes required by businesses or researchers.

6. Acknowledgements

The authors express their gratitude to the Department of Computer Science, University of Technology—Iraq and the staff for providing the software support and some references related to this work.

7. Conflict of interest

The authors declare that they have no conflicts of interest.

References

[1] S., Dash, S.K., Shakyawar, M., Sharma and S., Kaushik. "Big data in healthcare: management, analysis and future prospects," *Journal of Big Data*, vol. 6, no. 1, pp.1-25, 2019.

[2] B. M. Balachandran and S. Prasad, "Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence," *Procedia Computer Science*, vol. 112, pp. 1112–1122, 2017.

[3] C. Yadav, S. Wang, and M. Kumar, "Algorithm and approaches to handle large Data- A Survey," *International Journal of Computer Science and Network*, vol. 2, no. 3, 2013.

[4] J. R. Naif, G. H. Abdul-Majeed, and A. K. Farhan, "Internet of Things Security using New Chaotic System and Lightweight AES," *Journal of Al-Qadisiyah for computer science and mathematics*, vol. 11, no. 2, pp. 45–52, 2019.

[5] S. Smys and J. S. Raj, "Internet Of Things And Big Data Analytics For Health Care With Cloud Computing," *Journal of Information Technology*, vol. 1, no. 01, pp. 9–18, 2019.

[6] A. Agrawal and A. Choudhary, "Health Services Data: Big Data Analytics for Deriving Predictive Healthcare Insights," *Data and Measures in Health Services Research*, pp. 1–17, 2016.

[7] S., Al Mayahi, A., Al-Badi and A. Tarhini, "Exploring the potential benefits of big data analytics in providing smart healthcare," *International Conference for Emerging Technologies in Computing*, Springer, pp. 247-258, 2018.

[8] H. Fang, Z. Zhang, C. J. Wang, M. Daneshmand, C. Wang, and H. Wang, "*A survey of big data research*," *IEEE Network*, vol. 29, no. 5, pp. 6–9, 2015.

[9] C. M. Olszak and M. Mach-Król, "A conceptual framework for assessing an organization's readiness to adopt big data," *Sustainability*, vol. 10, no. 10, p. 3734, 2018.

[10] K. Batko and A. Ślęzak, "The use of Big Data Analytics in healthcare," *Journal of Big Data*, vol. 9, no. 1, pp. 1-24, 2022.

[11] L. Wang and C. A. Alexander, "Applications of Automated Identification Technology in EHR/EMR," *International Journal of Public Health Science (IJPHS*), vol. 2, no. 3, pp. 109–122, 2013.

[12] S. S. Abdul-Jabbar, L. E. George, "Fast Text Analysis Using Symbol Enumeration and Hashing Methodology," *Iraqi Journal of Science*, vol. 58, no. 1B, pp. 345–354, 2017.

[13] A. Jindal, A. Dua, N. Kumar, A. K. Das, A. V. Vasilakos, and J. J. Rodrigues, "Providing Healthcare-as-a-Service Using Fuzzy Rule Based Big Data Analytics in Cloud Computing," *IEEE Journal of Biomedical and Health informatics*, vol. 22, no. 5, pp. 1605–1618, 2018.

[14] M. T. Tamam, A. J. Taufiq, A. Kusumawati, and D. Setiawan, "Visualization of measurement of heart rate, oxygen saturation in blood and body temperature with a non-invasive method," *IOP Conference Series: Materials Science and Engineering*, vol. 674, no. 1, 2019.

[15] H. A. Alatabi, and A. R. Abbas, "Sentiment Analysis in Social Media using Machine Learning Techniques," *Iraqi Journal of Science*, Vol. 61, no. 1, pp. 193–201, 2020.

[16] D. Boyd and K. Crawford, "Six provocations for big data," *A decade in internet time: Symposium on the dynamics of the Internet and society*, pp. 1–17, Sept. 2011. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431

[17] F. Hassan and S. F. Behadili, "Modeling Social Networks using Data Mining Approaches-Review," *Iraqi Journal of Science*, vol. 63, no. 3, pp. 1313–1338, 2022.

[18] A. K. Bhadani and D. Jothimani, "Big data: Challenges, opportunities, and realities," *Effective big data management and opportunities for implementation*, IGI global, pp. 1–24, 2016.

[19] R. Sajjad and A.R. Dhary, "Using big data in telecommunication companies: A case study," *African Journal of Business Management*, vol. 14, no. 7, pp. 209–216, 2020.

[20] K.S. Kaswan, J.S. Dhatterwal, H. Sharma, and K. Sood, "Big Data in Insurance Innovation," *merald Studies in Finance, Insurance, and Risk Management*, Emerald Publishing Limited, Bingley, pp. 117-136, 2022.

[21] L. Zheng, and L. Guo. "Application of Big Data Technology in Insurance Innovation." International Conference on Education, *Economics and Information Management*, pp. 285-294. Atlantis Press, 2020.

[22] U. Srivastava and S. Gopalkrishnan, "Impact of big data analytics on banking sector: Learning for Indian Banks," *Procedia Computer Science*, vol. 50, pp. 643–652, 2015.

[23] V. Chang, L. Xiao, Q. Xu, and M. Arami, "A Review Paper on the Application of Big Data by Banking Institutions and Related Ethical Issues and Responses," *2nd International Conference on Finance, Economics, Management and IT Business*, SciTePress, pp. 115–121, 2020.

[24] J. Lawler and A. Joseph, "Big Data Analytics Methodology in the Financial Industry," I*nformation Systems Education Journal*, vol. 15, no. 4, pp. 38–51, 2017.

[25] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decision support systems*, vol. 50, no. 3, pp. 559–569, 2011.

[26] M. S. Delgosha, N. Hajiheydari, and S. M. Fahimi, "Elucidation of Big Data Analytics in Banking: a Four-stage Delphi Study," *Journal of Enterprise Information Management*, vol. 34, no. 6, pp. 1577–1596, 2021.

[27] M. A. Islam, F. B. A. Kasem, S. U. Z. Khan, T. Habib, and F. Ahmed, "Cloud Computing in Education: Potentials and Challenges for Bangladesh," *International Journal of Computer Science, Engineering and Applications*, vol. 7, no. 5, pp. 11–21, 2017.

[28] J. E. Beck and J. Mostow, "How who should practice: Using learning decomposition to evaluate the efficacy of different types of practice for different types of students," In *International conference on intelligent tutoring systems*, Springer, Berlin, Heidelberg, vol. 5091, pp. 353–362, 2008.

[29] D. Waga, E. Makori, and K. Rabah, "Utilization of Cloud Computing in Education and Research to the Attainment of Millennium Development Goals and Vision 2030 in Kenya," *universal Journal of Educational Research*, vol. 2, no. 2, pp. 193–199, 2014.

[30] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: Current state and future opportunities," *Proceedings of the 14th international conference on extending database technology*, pp. 530–533, 2011.

[31] I. Pawełoszek and J. Wieczorkowski, "Big data as a business opportunity: An educational perspective," *2015 Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE*, vol. 5, pp. 1563–1568, 2015.

[32] D. Jelonek, "Big Data Analytics in the Management of Business," *MATEC Web of Conferences*, vol. 125, pp. 1–6, 2017.

[33] A. M. Ali and A. K. Farhan, "Enhancement of QR Code Capacity by Encrypted Lossless Compression Technology for Verification of Secure E-Document," *IEEE Access*, vol. 8, pp. 27448–27458, 2020.

[34] T. Ercan, "Effective use of cloud computing in educational institutions," Procedia-Social and Behavioral Sciences, vol. 2, no. 2, pp. 938–942, 2010.

[35] C. T. S. Xue and F. T. W. Xin, "Benefits and Challenges of the Adoption of Cloud Computing in Business," International Journal on Cloud Computing: Services and Architecture, vol. 6, no. 6, pp. 01–15, 2016.

[36] A. Rashid and A. Chaturvedi, "Cloud Computing Characteristics and Services A Brief Review," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 2, pp. 421–426, 2019.

[37] R. Choubey, R. Dubey, and J. Bhattacharjee, "A Survey on Cloud Computing Security, Challenges and Threats," *International Journal on Computer Science and Engineering (IJCSE)*, vol. 3, no. 3, pp. 1227–1231, 2011.

[38] D. Zissis and D. Lekkas, "Addressing cloud computing security issues," *Future Generation computer systems*, vol. 28, no. 3, pp. 583–592, 2012.

[39] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and Good practices," *In 2013 Sixth international conference on contemporary computing (IC3)*, IEEE, pp. 404–409, 2013.

[40] G. H. Abdul Majeed, A. Kadhim, and R. S. Ali, "Retrieving Encrypted Query from Encrypted Database Depending on Symmetric Encrypted Cipher System Method," *Diyala Journal for Pure Sciences*, vol. 13, no. 1, pp. 174–198, 2017.

[41] Z. K. Mahyob, R. A. Ogla, and S. M. Zeki , "The Techniques of Based Internet Key Exchange (IKE) Protocol to Secure Key Negotiation," *IRAQI JOURNAL OF COMPUTERS, COMMUNICATIONS, CONTROL AND SYSTEMS ENGINEERING*, vol. 22, no. 3,147-154, 2022.

[42] S. Yu, "Big Privacy: Challenges and Opportunities of Privacy Study in the Age of Big Data," *IEEE Access*, vol. 4, pp. 2751–2763, 2016.

[43] S. A. Jassim, and A. K. Farhan, ," Designing a New Lightweight AES Algorithm to Improve the Security of the IoT Environment," *IRAQI JOURNAL OF COMPUTERS, COMMUNICATIONS, CONTROL AND SYSTEMS ENGINEERING*, vol. 22, no. 2, 96-108, 2022.

[44] B. Agrawal, "Scalable Data Processing and Analytical Approach for Big Data Cloud Platform", *PhD Thesis*, University of Stavanger, 2017.

[45] D. Sharma, G. Pabby, and N. Kumar, "Challenges Involved in Big Data Processing & Methods to Solve Big Data Processing Problems," *IJRASET*, vol. 5, no. 8, pp. 841–844, 2017.

[46] M. Saadoon, S. H. A. Hamid, H. Sofian, H. H. Altarturi, Z. H. Azizul, and N. Nasuha, "Fault tolerance in big data storage and processing systems: A review on challenges and solutions," *Ain Shams Engineering Journal*, vol. 13, no. 2, p. 101538, 2022.

[47] S. H. Sabah, and M. S. Croock, Software Engineering Based Fault Tolerance Method for Wireless Sensor Network. *IRAQI JOURNAL OF COMPUTERS, COMMUNICATIONS, CONTROL AND SYSTEMS ENGINEERING*, vol. 20, no. 4, pp. 21-28, 2020.

[48] A. Labrinidis and H. V Jagadish, "Challenges and Opportunities with Big Data," *Proceedings of the VLDB Endowment*,vol.5, no. 12, pp. 2032-2033, 2012.

[49] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Information, communication & society*, vol. 15, no. 5, pp. 662–679, 2012.

[50] B. kalra, S. Yadav, and D.K. Chauhan "A Review of Issues and Challenges with Big Data," *International Journal of Computer Science and Information Technology Research*, vol. 2, no. 4, pp. 97-101, 2014.

[51] L. Zhang, "Big Data Analytics for Fault Detection and its Application in Maintenance, " *PhD dissertation*, Luleå University of Technology, 2016.

[52] J. Manyika, M. Chui., B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Hung Byers, "Big data: The next frontier for innovation, competition, and productivity". *McKinsey Global Institute,* 2011.

[53] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Networks Application*, vol. 19, no. 2, pp. 171–209, 2014.

[54] M. Hossain, "Dynamic scaling of a web-based application in a Cloud Architecture," 2014. Available: https://www.diva-portal.org/smash/record.jsf?pid=diva2:699823

[55] J. Sen, "Security and privacy issues in cloud computing," *Cloud technology: concepts, methodologies, tools, and applications*, pp. 1585-1630. IGI global, 2015.

[56] D. Tafani, B. Kantarci, H.T. Mouftah, C. McArdle, and L.P. Barry, "Towards energy efficiency for cloud computing services". *Communication Infrastructures for Cloud Computing*, IGI Global, pp. 306-328, 2014.

[57] B. Berisha, E. Mëziu and I. Shabani, "Big data analytics in Cloud computing: an overview," *Journal of Cloud Computing*, vol. 11, no. 1, pp.1-10, 2022.

[58] C. Lee, Z. Luo, K.Y. Ngiam, M. Zhang, K., Zheng, G. Chen, B.C. Ooi, and W.L.J. Yip, "Big healthcare data analytics: Challenges and applications," *Handbook of large-scale distributed computing in smart healthcare*, pp.11-41, 2017.