EJSIT
European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 3 | Number 3 | 2023

Comparative Analysis of Blockchain and Database

Joseph Diema Enoch[1*], Bourdillon O. Omijeh[2] and Remigius Obinna Okeke[3]
[1]Department of Computer Engineering, Rivers State University, Port Harcourt, Nigeria
[2,3]Department of Electrical/Electronic Engineering, University of Port Harcourt, Nigeria

ABSTRACT

The purpose of this research is to perform an in-depth comparative analysis on blockchain, centralized and distributed database management system so as to aid researchers ascertain the correct application of these technologies. Most practitioners are already familiar with centralized database system which is commonly used for general purposes. The findings in this study with respect to the features of blockchain and database were put together to develop a program flowchart that provides the sequence of steps to guide researchers in making the right decision for adopting any of the listed solutions. Furthermore, it was noted that blockchain has gained much attention in most sectors including the educational sector. From the critical review of over 200 articles in Google Scholar, the study was able to analyze ten key factors and drew inference from it to produce the table that summaries the key comparison factor for adopting blockchain or database. Finally, the resolve from the flowchart and the table could serve as a guide to any researcher interested in deploying any of these technologies to quickly make the right decision on when to use blockchain, centralized or distributed database management system.

Keywords: Analysis, Blockchain, Comparative, Centralized and Distributed Database

INTRODUCTION

Generally human or researchers search for technology they could deploy to provide them with the required ingredients they need in developing a robust solution. It is most likely that many researchers use less than 50% of the capabilities of some of the technologies they deploy to solve a problem. Moreover, at often times, they deploy an inappropriate technology that could still do the task. The essence of this study is to investigate the strength, weakness, advantages or disadvantages of deploying blockchain or database management system. The word blockchain, comprises of two key words, namely block and chain. Where all the data and transactions are stored in a blockchain are viewed as blocks. These blocks are linked together with the sha256 hashing algorithm of the previous and the current block hash. It can be defined as a ledger technology that is decentralized and distributed, which records the attribution of an asset in digital format (Builtin, 2022). It is a promising and revolutionary technology due to its potential of reducing risk, eliminates scam and builds trust and transparency in an adequate manner for several purposes. Sophisticated innovations and complex mathematical algorithm that are very difficult to manipulate or hack by fraudsters (Orcutt, 2018). The centralized database is the most commonly used database management system for most business solutions. It is a database system which is located and stored in a single server connected to a network or standalone. Single or more users finds it easier to access and get a total view of the record since it is localized on a server. It's also simpler to manage, update and back up information. A distributed database management system is not prone to failure like the centralized database that has a single point of failure. It is a database system in which the database is replicated on multiple servers connected on the same network or a different network linked via the internet. Single or multi-users could access the closest database server, which means it offers a faster way of retrieving data. When there is a failure

*Corresponding Author

on one database server, users could still gain access to data through the other servers on the network (Tech Insights for Professionals, 2023).

Blockchain technology provides some key features of which its distributed nature allows everyone on the decentralized blockchain network to access to transactions and view the entire system in a more transparent way. Another key feature is its standardized rules which requires that for every transaction processes the same sets of rules are applied. In privacy each participant identity is hidden because blockchain creates an address for all users. It is this address that is revealed in the user transactions. Especially in vote cast, the real identity of the user is not revealed. Furthermore, its audit-ability makes it easy for records to be traced or confirm its validation as transactions are being made. Most importantly its security provides a powerful cryptography that makes transaction and data very secured for each user (Angeles, 2018; Monrat *et al.*, 2019; Shi *et al.*, 2020; Zheng *et al.*, 2018). The centralized database system consist of a central server and many clients. The server communicate and responds to the clients requests. If the server fails, the system will be down. Whereas, the distributed database management system (DDBMS) is a network of distributed database servers. Each server has a copy of the database. An update in any of the database triggers the automatic update of all the other databases on the network. The DDBMS could be autonomous in this case each database server function independently or non-autonomous in which a master database system coordinates data update across the homogeneous nodes (Enoch, 2023).

REVIEW OF RELATED RESEARCH

Several studies have been carried out on the critical analysis of blockchain versus database. A study carried out by Ruan *et al*. (2021) compared blockchains and distributed database (dichotomy and fusion) in terms of security and performance. The study also developed a mechanism in filling the gap in the new trend of merging blockchain and distributed database. It further proposed a taxonomy that illustrates the dichotomy across four dimensions, namely replication, concurrency, storage, and sharding. It discussed how the design choices are driven by two goals: security for blockchains, and performance for distributed databases. The study carried out by Enoch, (2021) performed a concise review of cutting edge technologies to aid researchers ascertain how these technologies works especially the use of blockchain technology. Bergman *et al*. (2020) focused on the performance of permissioned blockchains and distributed databases. The research investigate the insert and read latency of Hyperledger Fabric to Apache Cassandra in four experiments. Latency measurements were performed, for small networks, the insert latency of Cassandra is twice as high as that of Fabric, whereas for larger networks Fabric has almost twice as high insert latencies as Cassandra. Fabric has around 40 ms latency for reading data and Cassandra between 150 ms to 250 ms, thus it scales better for reading.

A study carried out by Liu *et al.* (2020) proposed the use of blockchain and distributed ledger-based improved biomedical security system (BDL-IBS) to improve security and privacy of data across healthcare applications. The results of the study showed that blockchain-based solution is seamless, easier and faster in enhancing data security and privacy. Wang *et al*. (2020) researched and analyzed the distributed database of blockchain and non-blockchain. The study compared the performance of BigchainDB and two distributed databases: HadoopDB and Hive. The experiments in the study show that BigchainDB was better than HadoopDB and Hive in transaction creation, query ot fuzzy query and also in reading and writing. Chen *et al*. (2018) conducted a comparative testing on performance of blockchain and relational database. Using Ethereum and MySQL, the results showed that single transaction on blockchain network was 1:10 of MySQL and also the time spent in processing single transaction, was 4:100 times. The study of Chowdhury *et al*.

(2018) presented a critical analysis of blockchain versus database. Based on this analysis, the study develop a decision tree diagram that will aid practitioners to choose the appropriate technology for their use cases.

The detailed examination conducted in this study showed that blockchain and distributed database should not be used for general business solution like centralized database (Enoch, 2023). The reviews also reveal that not much work have been done in the comparison of blockchain and database in terms of scalability, traceability, automation and other key factors in order to guide researcher in making the right decision on the technology to deploy, which this research seeks to accomplish.

METHODOLOGY

Comparative research approach was used for this study. The analysis was based on ten key factors that would determine blockchain or database usability. This research as stated earlier carried out a comparative analysis between database and blockchain using different criteria described in (Chowdhury *et al.*, 2018) and the adaptation of blockchain were considered. This research also provided more in-depth analysis compared to the study of Chowdhury *et al.* (2018) by including the instant traceability, audit-ability and automation features of blockchain and scalability features of database.

Analysis based on Security of Database versus Blockchain

The management and maintenance of data in a conventional database system is implemented either centrally or on a distributed bases. In the traditional database system, the deployed mechanism for enforcing restrictions and security of data is the access control method. A central database system is easily prone to attacks and it is vulnerable, especially when the database administrator is compromised or hacked. This is similar for a distributed database system. The only difference is that in distributed database system data are spited across multiple physical servers, and often across more than one data centre. A compromise in one data centre would be replicated in other data centre. Moreover, staff who possess legal right could modify data for personal gain or interest. However, on the contrary, blockchain technology was developed with the adaptation of high security measures with the use of a hashing algorithm and also a consensus system. The consensus convergence strength is determined by the number of nodes or users in the system. One of the rules that governs the use of blockchain demands that a new block can only be added when a minimum of 51% of the nodes are in consensus. This implies that it is only when 51% of the node are compromised or miss manage by compromised users that it can accept an invalid transaction as a valid data. This security features of blockchain makes it almost impossible to modify data.

Analysis based on Trust

The study found out that one of the key features of blockchain is its immutability capability. The consensus algorithm of the decentralized blockchain network guarantees the ability of immutability. The acceptance or validity of data or transaction is determined by all the participating nodes via the consensus algorithm. It is worthwhile to note that the access level and capability of all the participating nodes is the same. This is a feature in blockchain that promotes confidence and therefore, built trust in the system. It proofs that the system is democratic in nature. Whereas, in a distributed or centralize database management system, the control is central or distributed to a few. The administrator(s) decides the privileges assigned to users, they control the management of the system. This implies that the system can only be trusted if the database administrator is walking in integrity or can be trusted. In addition, if the system is not connected to a public network with efficient firewalls.

Analysis based on Confidentiality and Privacy

The data in a blockchain are not stored encrypted, even though a lot of people think so. It is a misconception people have about blockchain solution. The contents in a blockchain are not all encrypted however they are digitally signed. Each block is mined with a previous hash containing the transactions of the previous block and the current block hash containing the present block transactions together with the previous hash data. The fact about blockchain is that it is a ledger system that is public. It allows anyone to participate in the verification of transactions on the decentralized blockchain network. The participating users' confidentiality or privacy are protected by using a public cryptography key. However, it reveals the data in the transactions and the participating users, for instance in crypto-currency the amount of coin received or transferred. There are proposals by researchers for strong anonymization using cryptographic means such as Zero Knowledge protocol (Fortnow, 1987; Chowdhury *et al.*, 2018). In the traditional database system access control method can by deploy to define users access privileges to enforce confidentiality.

Analysis based on Traceability

The activities or transactions that is carried out by any user can be documented from the origin or source in blockchain. It allows you to develop an audit system that records the origin of every event or action within the system. Mostly, in the university community where users are concerned about mutilation issues surrounding academic results or organization worried by forgery, fraud, fake, or corrupt records. Blockchain solution guarantee the traceability and can provide the evidence of the transaction origin or source. Whereas in a centralized or distributed database system it is difficult to trace the origin when data is modified by a compromised administrator or hacker.

Analysis based on Automation

Program instructions can be created and stored in a blockchain to execute when some predefined criteria are fulfilled. This predetermined sets of program instruction is implemented as start contracts. It is a reliable means of automating processes and transaction in blockchain with little or no human intervention. In addition, it does not require the participation of any third party to verify it since the term of the contract is already defined and it is being followed by the program. Hence, it is a faster, reliable, efficient and dependable solution for any automation system that require accuracy and un-mutilated responses. Meanwhile, in the traditional database system, the automation reliability depend mainly on the database administrator or developer's accountability. Most importantly, requires third party verification of the contract.

Analysis based on Robustness/Fault Tolerance

Robustness and fault tolerance is achieved by using the distributed computing nature of the decentralized blockchain network. In a blockchain solution, transaction data are kept saved in a distributed manner. Copies of the whole chain are available in all the participating or connected node on the decentralized blockchain network. Hence cyber attacks that affects the traditional database system can easily be overcome by this system. The advantage distributed database system have over centralized database system is that it can survive denial of service (DoS) attacks since failure on one server does not affect the entire system. However, it is vulnerable to distributed denial of service (DDoS) attack. Blockchain solution make DoS and DDoS attacks unachievable on the blockchain network. Moreover, the compromise or failure of a node or some nodes does not affect blockchain performance on the network.

Analysis based on Performance

The Ethereum blockchain transactions could take an average of 15 seconds to 5 minutes to process, whereas, bitcoin is extremely slow. The confirmation of transaction on the network could take about 10 minutes. Depending on the situation, the time could extend to about 60 minutes especially, when there is soft fork in the network (Lin & Liao, 2017). The speed of traditional database system is tremendous. It can manage thousands of data transaction in a second. A good example is the banking sector, where electronic payment transaction from Master-Card, Verve-Card or Visa-Card system can manage about 50,000 transaction in a second. In this system it is possible for the administrator to increase the bandwidth to allow more transactions when the network of transactions increases. Presently, there are ongoing research to improve on the performance of blockchain based on the consensus method used in blockchain network. Consensus algorithm like Ethash (Mukhopadhyay *et al.*, 2016) and X13 (Rabah, 2017) with about 10 to 20 seconds could achieve consensus. Moreover, there are two faster consensus algorithm used in permission blockchain with low energy consumption, they are Practical Byzantine Fault Tolerance (PBFT) algorithm and Istanbul Byzantine Fault Tolerance (IBFT) algorithm (Honnavalli *et al*., 2020).

Analysis based on Redundancy

The distributed nature of blockchain make it easy to eliminate redundancies in the system. Copies of the whole chain are available in all the participating or connected node on the decentralized blockchain network. The update is done on a continuous bases as transactions are made. Hence, it does not have the problems associated with a single point of failure. This feature of blockchain is also replicated in distributed database system. In blockchain technology, a new blocks are added to all the nodes on the blockchain network at the same time to ensure synchronization. Whereas, in a central database system it has a single point of failure that could increase the level of redundancy or down time. It is important to state that in a centralized database system, only the central system has copies of all transaction. However, in a distributed database system, copies of all transactions are forwarded to other servers in a distributed network. An update in the database triggers the automatic synchronization of the update to other database server in the distributed system.

Analysis based on Auditability

Blockchain generates an audit system for every transaction and records the source of that data. In the academic community, when people are worried about mutilation of academic data, or in a situation where a problem with fraud, forgery, and fake results exists. Blockchain solution could assists in providing tractability with evidence from the stored audit records. Information like timestamp, transaction id from the audit record can be used to verify and trace transaction of previous records. Thereby making the information saved in the blockchain transparent and traceable. Blockchain based hashing algorithm are used to store the records in a chronological manner in the distributed ledger and the availability of this irreversible records to all network users. Whereas in traditional database management system the audit ability of the system depends on trusted parties or a database administrator with integrity. The audit records can be modified if the system is compromised and hence not traceable to the user. To solve this issue is to combine the database system with a blockchain solution to store the audit information.

The Adoption of Blockchain

Sequel to the analyzed ten key factors, it is evident that the application of blockchain solution is not for the general development of any business solution like database

management system. However, it should be wisely applied to maximize its key benefits and potentials. The developed program flowchart in Figure 1 clearly defines the steps or factors to consider in order to select the suitable technology to adopt before designing or implementing any business solution. It also illustrates the sequence of decisions that would guide the system architect or business/system analyst to make clear decision on when to use and not to use database or blockchain. Most importantly, blockchain technology usage is appropriate to deploy in the cases where different groups of common interest or goals are participating in a joint venture. It is most likely that trust would be an issue of utmost concern among them. Blockchain has gained popularity in the area of SCM (supply chain management) from our review. It is a management system where the flow of data, goods and services from the procurement stage to the delivery stages is managed by different groups. Another typical example were this technology could be used or be beneficial is in the academics. In the area of recommendation and approval of students' results by majority or all the members of senate after due consultation and consideration applying the consensus mechanism of blockchain. In addition, it can be deployed in the health sector to automate prompt responses to critical health situation and treatment. Finally, group of organizations working on a government established project as a consortium would require a trusted system to manage transactions between them. In the cases where you have trusted third party, like the bank mediating between a supplier and a distributor. The distributor and the supplier can walk together because they all trust the bank. In addition, in a case where performance and scalability are required it is advisable to use the traditional database system. However, most of the times, it is most likely that obtaining a trusted third parties could be very difficult and also risky. Therefore, it would be of great benefit to adopt blockchain especially if most of the listed conditions are met in Figure 1. There are two major ways blockchain can be implemented either as a public blockchain or a private blockchain. The public blockchain is used when the data or transaction stored in the blockchain requires the access and verification of the public while private blockchain is best when the data transactions is specific to an organization with a simplified and faster consensus algorithm. It is worthwhile to note that if the traceability of origin of data transaction and storage of large volume of data is required. It is recommended to use off chain storage blockchain whereas if data integrity is of great importance, then use on chain storage blockchain.

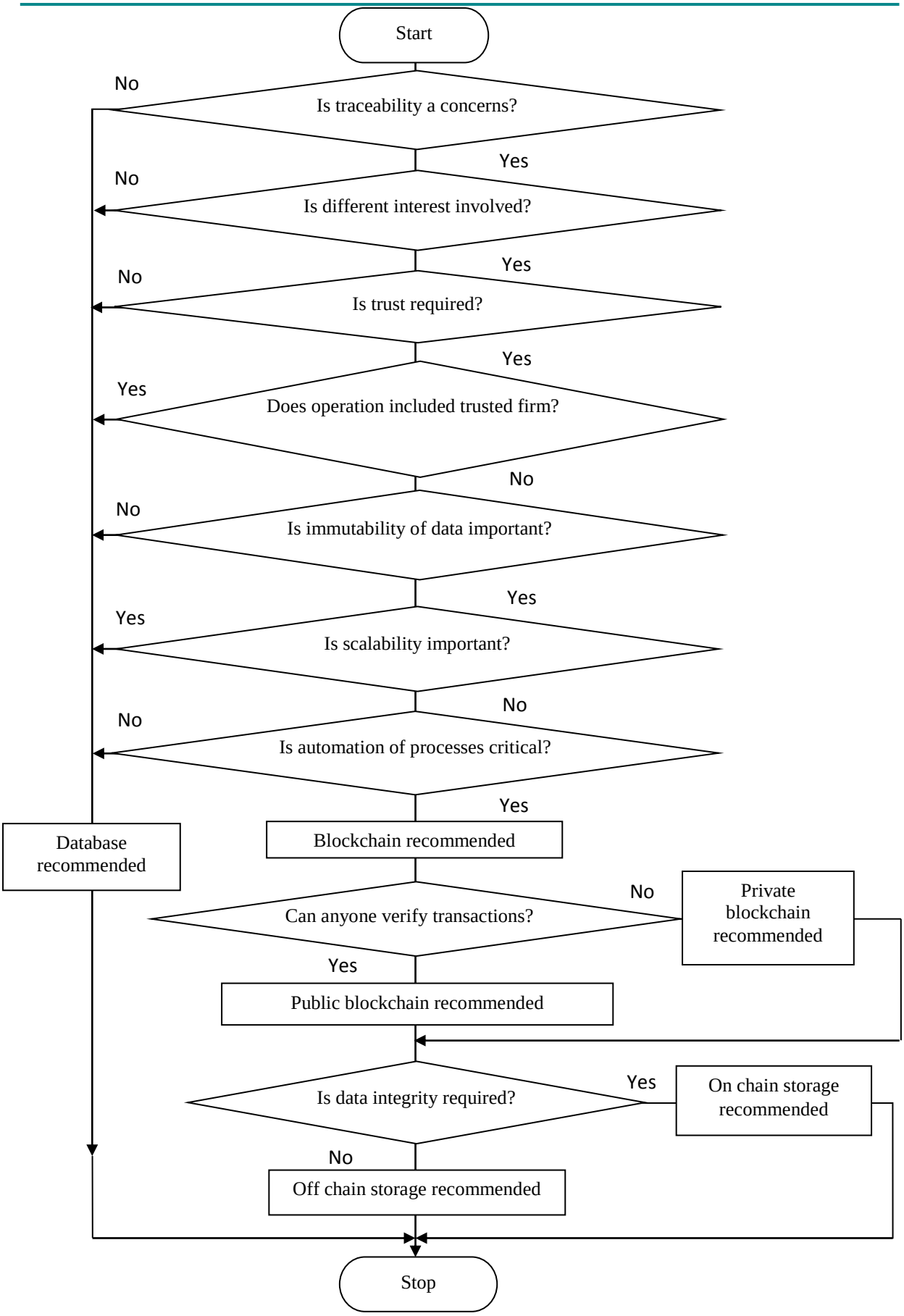

Figure 1: Blockchain Adoption Flowchart

RESULTS AND DISCUSSION

The findings and the results from the critical analysis of the comparative study were discussed in this section. The research also considered the applicability, opportunities and restrictions of blockchain in the educational sector among others. Having evaluated and narrowed the research articles to 200 research work from a pull of reviewed journal articles in Enoch (2023) PhD thesis surveyed from Google scholar. A guiding model was prescribed in the developed program flowchart in Figure 1. The study findings also showed that majority of the 200 articles were either theoretical or empirical research relating to blockchain technology. The research went further to identify 80 papers that were related to blockchain implementation. In comparison, it was found that among these 80 articles, the unique and special features of blockchain were used appropriately by 50 of them. Meanwhile, the other 30 articles applied the traditional databases system to achieve similar functionality. Furthermore, it was found that 30 papers out of 50 papers were related to Health, Finance and SCM (supply-chain-management) system and the rest were on educational system improvement. The research investigation also shows that blockchain technology has become known as a valuable and productive solution with the following instances: it was applied, at the Massachusetts Institute of Technology (MIT) to protect and validate the certificates that it issued. Secondly, Sony Global Education, Tokyo, Japan, used it to secure data on its trainees' competencies and productivity. Thirdly, University of Nicosia, Cyprus, were first to use smart contracts and accept cryptocurrency as a form of payment and finally, universities in & outside of Russia used it to improve their educational system. The review further indicates its relevance and sustainability in education. However, there were areas the technology was used inappropriately. One of them is in smart home scenarios, since it was built on a personal network in which trust is not a major concern. Therefore, it was not appropriate to use blockchain.

Table 1: Result of the adoption of Blockchain versus Database

Comparison Factor	Blockchain	Central/Distributed Database	Recommendation
Security	It is developed with cryptographic by default	Traditional access control method is used	Blockchain
Trust	No trusted party is required for it to function	A trusted party is required	Blockchain
Traceability	It documents the origin of data in every transaction	It need a trusted documentation provider	Blockchain
Automation	Transactions can be automated with some predefined program codes	It need trusted automation support	Blockchain
Auditability	It keeps an irreversible record of all transaction sources	It need a trusted documentation provider	Blockchain
Fault Tolerance /Robustness	Distribution of transactions are done at the same time among nodes	Data updates, triggers update on other servers in a distributed system or stored in database.	Blockchain or Distributed Database
Redundancy	Copies of the entire chain is in all the participating node with the latest copy by default	Central system has a copy but distributed system has copies on all the servers	Blockchain or Distributed Database

Confidentiality of Data	Data are visible to all nodes by default	Authorized person have restricted access privileges	Database
Performance	Consensus require some time to achieve (e.g. Bitcoin takes 10 to 60mins)	Immediate execution and implementation of data update, it is faster	Database
Scalability	Require supports from other system e.g., cloud storage	Perform more efficiently	Database

CONCLUSION

The comparative analysis among blockchain, distributed and centralized database management system were carried out successfully. The research findings have shown that blockchain solution had gained much attention from most sectors including the educational sector. The technology have been used to solve different kinds of problems, most especially in the finance sector. The survey in this research had identified some areas in the educational sector that have recorded some feat in the university community especially in the area of stopping result falsification. In addition, the study also analyzed centralized and distributed database management system. From the outcome of the study, it can be deduced that blockchain would be the best recommended option, if traceability, robustness, trust, automation, Auditability, redundancy and security of data are of great concerns or importance. However, centralized database system should be deployed if performance, scalability and confidentiality, are the major concerns. In addition, if performance, scalability, redundancy, robustness and confidentiality are the requirement of the system then distributed database management system should be used. Finally, the findings from the critical analysis of the study were used to develop a program flowchart and a table that summarizes the key features that would guide researchers ascertain the correct application of blockchain with respect to the traditional database management system. This study hopes to guide and direct researchers the appropriate way to use blockchain technology. It would help them make the right decision in order to avoid misuse or unsuccessful implementation of blockchain solution.

REFERENCES

Angeles, R. (2018). Blockchain-Based Healthcare: Three Successful Proof-of-Concept Pilots Worth Considering. *Journal of International Technology and Information Management*, *27*(3), 47-83.

Bergman, S., Asplund, M., & Nadjm-Tehrani, S. (2020). Permissioned blockchains and distributed databases: A performance study. *Concurrency and Computation: Practice and Experience*, *32*(12), e5227.

Builtin (2022). Blockchain: What Is Blockchain Technology? How Does It Work? Retrieved from https://builtin.com/blockchain**.**

Chen, S., Zhang, J., Shi, R., Yan, J., & Ke, Q. (2018, July). A comparative testing on performance of blockchain and relational database: Foundation for applying smart technology into current business systems. In *International Conference on Distributed, Ambient, and Pervasive Interactions* (pp. 21-34). Springer, Cham.

Chowdhury, M. J. M., Colman, A., Kabir, M. A., Han, J., & Sarda, P. (2018, August). Blockchain versus database: a critical analysis. In *2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th*

IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE) (pp. 1348-1353). IEEE.

Enoch, J. D. (2021). Emerging Information Technologies in the 21st Century. *Proceedings of the International Conference on Newviews in Engineering and Technology* (Maiden Edition, pp. 207-216). Faculty of Engineering, Rivers State University, Port Harcourt. Nigeria, 27th October 2021, ICNET2021-021.

Enoch, J. D. (2023). *Development and Implementation of Blockchain Solution to Enable Data Security in University of Port Harcourt* [Ph.D. Thesis]. Centre for Information and Telecommunication Engineering, University of Port Harcourt. Rivers State, Nigeria.

Fortnow, L. (1987, January). The complexity of perfect zero-knowledge. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing* (pp. 204-209).

Honnavalli, P. B., Cholin, A. S., Pai, A., Anekal, A. D., & Anekal, A. D. (2020). A study on recent trends of consensus algorithms for private blockchain network. In *Blockchain and Applications: 2nd International Congress* (pp. 31-41). Springer International Publishing.

Liu, H., Crespo, R. G., & Martínez, O. S. (2020, September). Enhancing privacy and data security across healthcare applications using blockchain and distributed ledger concepts. In *Healthcare*, *8*(3), 243. Multidisciplinary Digital Publishing Institute.

Lin, I. C., & Liao, T. C. (2017). A survey of blockchain security issues and challenges. *Int. J. Network Security*, *19*(5), 653-659.

Monrat, A. A., Schelén, O., & Andersson, K. (2019). A survey of blockchain from the perspectives of applications, challenges, and opportunities. *IEEE Access*, *7*, 117134-117151.

Mukhopadhyay, U., Skjellum, A., Hambolu, O., Oakley, J., Yu, L., & Brooks, R. (2016, December). A brief survey of cryptocurrency systems. In *2016 14th annual conference on privacy, security and trust (PST)* (pp. 745-752). IEEE.

Orcutt, M. (2018). MIT technology review: how secure is blockchain really? It turns out "secure" is a funny word to pin down. Retrieved from https://www.technologyreview.com/2018/04/25/143246/how-secure-is-blockchain-really/

Rabah, K. (2005). Theory and implementation of data encryption standard: A review. *Information Technology Journal*, *4*(4), 307-325.

Ruan, P., Dinh, T. T. A., Loghin, D., Zhang, M., Chen, G., Lin, Q., & Ooi, B. C. (2021, June). Blockchains vs. distributed databases: dichotomy and fusion. In *Proceedings of the 2021 International Conference on Management of Data* (pp. 1504-1517).

Shi, S., He, D., Li, L., Kumar, N., Khan, M. K., & Choo, K. K. R. (2020). Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. *Computers & security*, *97*, 101966.

Tech Insights for Professionals (Tuesday, March 7, 2023). Distributed vs Centralized: The Battle of the Databases. Retrieved from https://www.insightsforprofessionals.com/it/storage/distributed-vs-centralized-database#:~:text=A%20centralized%20database%20is%20one,information%20in%20a%20centralized%20database

Wang, Y., Hsieh, C. H., & Li, C. (2020, April). Research and analysis on the distributed database of blockchain and non-blockchain. In *2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)* (pp. 307-313). IEEE.

Zheng, Z., Xie, S. & Dai, H. (2018). Blockchain Challenges and Opportunities: A Survey. *Int. J. Web and Grid Services*, *14*(4), 352-375.