Journal of Theoretical and Applied Information Technology
15th September 2024. Vol.102. No. 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A SYSTEMATIC LITERATURE REVIEW ON BLOCKCHAIN FOR IMPLEMENTATION OF DATA GOVERNANCE FRAMEWORK

MUHAMMAD FEIZAL[1,2], ROSMAYATI MOHEMAD[2*], NOOR MAIZURA MOHAMAD NOOR[2]

[1]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, 15310 Banten, Indonesia
[2]Faculty of Computer Science and Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

Email: [1]dosen00318@unpam.ac.id, [2]rosmayati@umt.edu.my, [2]maizura@umt.edu.my

ABSTRACT

Data plays an increasingly pivotal role as it has the influence to shape the decision-making process of an institution, ultimately impacting its competitiveness and enhancing customer satisfaction. As technology continues to progress, the volume of data accessible is expanding from various diverse data sources, thereby leading to the possibility of data inaccuracies due to the significant challenges in integrating, maintaining consistency, and ensuring interoperability of the data. The accuracy of decision-making is greatly influenced by the trust level of the data source and the quality of the data. High data quality and a strong level of trust in data are essential in data governance to guarantee the accuracy, consistency, security, and accessibility of data. This, in turn, facilitates decision-making and resource management. Currently, the majority of public understanding regarding blockchain is that blockchain is the same as cryptocurrency. However, based on its characteristics, blockchain has the opportunity to be used for data management in various fields. The potential of blockchain apart from cryptocurrency in this case for data management is not yet widely understood by most people. The purpose of this paper is to summarize the development of blockchain role in supporting data governance using a systematic literature review (SLR) method through Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) and state-of-the-art. A total of 449 papers are collected from Scopus, IEEExplorer, Science Direct, and Google Scholar Search Engine. From these, 15 papers are selected. The findings indicate that blockchain technology has the potential to significantly improve data governance in a variety of areas, including data value assessment, data standardization and transformation, data transparency, data security, data sharing, and data storage. This can be done because blockchain has the capability to utilize cryptography and smart contract automatically. The key elements comprising blockchain-based data governance are users, access and integration protocol, and data storage.

Keywords: *Blockchain, Data Governance, Systematic Literature Review, SLR, Trust, Data Quality*

1. INTRODUCTION

The total amount of data created, captured, copied, and consumed globally is forecast to increase rapidly. Data is becoming crucial for decision making and resource management, which will influence competitiveness and increase customer satisfaction. Therefore, it is important to maintain the quality of the data. Apart from that, currently institutions are required to be able to collaborate with various other parties so that data sources can come from internal or external to the institution. The quality of the data can also influence the level of trust of the collaborating parties.

The volume of data generated, acquired, replicated and utilized is projected to persist in its rapid expansion. Recent forecasts indicate a significant increase in the global volume of data generated in the upcoming years. According to the International Data Corporation (IDC), the global data sphere will expand from 64.2 zettabytes in 2020 to over 180 zettabytes by 2025 [1]. The vast amount of data has

a significant influence on the quality of decision-making and resource management, leading to the formation of strategic plans, improved problem-solving, optimized operations, increased competitiveness, and enhanced customer satisfaction. Therefore, it is important to maintain the quality of the data. Apart from that, the collaboration and sharing of data among organizations can further enhance decision accuracy and efficiency, whilst optimizing resource utilization [2]. The quality of data also significantly influences the level of trust among collaborating parties, particularly within the context of data governance.

According to Data Management Association (DAMA) International, data governance is defined as the exercise of authority and control (planning, monitoring and enforcement) over the management of data assets [3]. Data governance provides direction and oversight for data management by establishing a system of decision rights over data. Data governance is required for consistency and balance between function in data management knowledge areas. In addition, data governance, which includes reference data management, data architecture, data security and metadata management, provides a foundation on which all other functions are dependent. Data governance program enables an organization to be data-driven, by putting in place the strategy and supporting principles, policies and stewardship practices that ensure the organization to get value from its data [4]. According to Data Governance Insitute, data governance is system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods [4]. One of the important things that needs to be ensured in building data governance is how data governance can guarantee high quality and reliable data. One technology that has the capability to record data and its changes is blockchain. The blockchain concept was put forward in a paper written by the pseudonymous person or persons who use the name Satoshi Nakamoto [5]. Blockchain is a digital database containing information (such as records of financial transactions) that can be simultaneously used and shared within a large decentralized, publicly accessible network and also the technology used to create such a database [6]. Blockchain is a mechanism for storing transaction records similar to a ledger, which can be accessed publicly. Where the recording of transactions is represented in interrelated blocks, and can grow when there are new blocks (transactions). To maintain the security and consistency of recording these transactions, blockchain utilizes asymmetric cryptography and a distributed consensus algorithm. [7]. The development of blockchain technology has gone through quite a long time. The history of blockchain development can be seen in Figure 1.

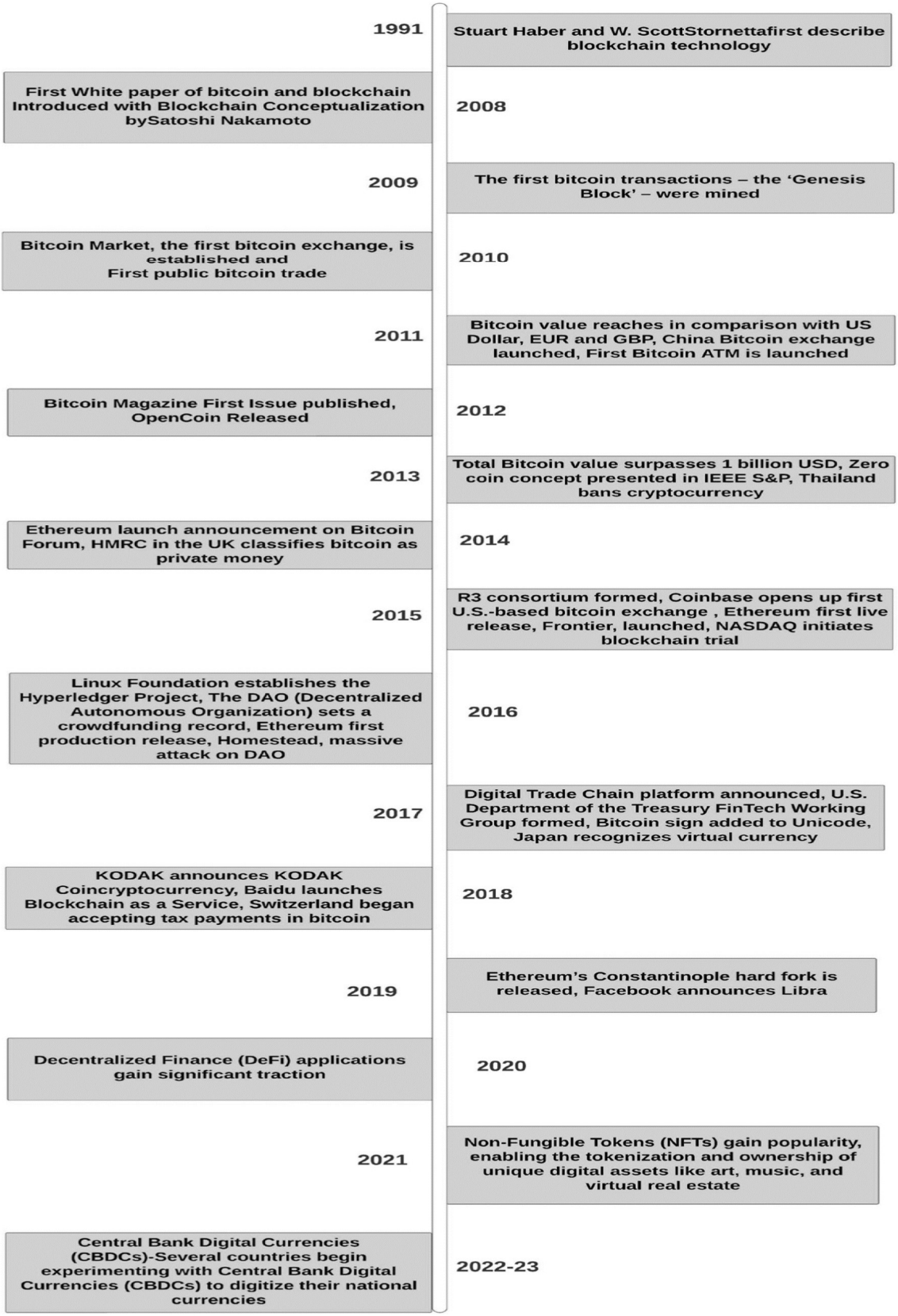


Figure 1: History of Blockchain Development [8]

Today's companies operate in an environment where data is produced in enormous quantities. Data has become critical to the success of companies. This led to a shift in strategy over two decades. Companies have the potential to cultivate a culture that uses data and business insights to drive all company decisions and become a data-driven company. Companies no longer base their decisions on intuition. Companies become data-driven company by encouraging stakeholders to use data as a basis for carrying out every business activity [9]. Businesses can benefit through optimizing data and formulating data-based insights, so that data becomes a valuable asset. At the same time, they are facing privacy and security law related to data such as General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), California Privacy Rights Act (CPRA) and Perlindungan Data Pribadi (PDP), as well as more complex consumer expectations around the use of their personal data. To face these challenges, especially in managing critical data, data-driven companies need to build collaboration between cross-functional teams. By automating data management, it can make it easier for business functions (such as finance, sales, marketing and customer support) in the company to access the data they need. This is in line with the increasing trend of self-service business models [10]. By looking at the characteristics of blockchain which allows the formation of a decentralized, audited, anonymous and persistent data recording system [7], there is great potential to utilize blockchain to overcome these challenges.

Blockchain has been widely used for cryptocurrency. But blockchain has wider potential than just cryptocurrency, including ensuring the implementation of data governance. However, there is no mapping of existing blockchain-based framework patterns and mapping of the role of blockchain for data governance. The objective to be achieved in this paper is to conduct a review on how developments in blockchain research can contribute to the implementation of automated data governance by conducting a systematic review of research related to this matter.

This paper is structured as follows. Section 2 explains the research methodology used. Includes an explanation of the criteria and research questions used. Section 3 discusses the results obtained from the analysis of existing literatures and the discussion about the results. Section 4 summerizes the conclusions of this paper.

2. RESEARCH METHODOLOGY

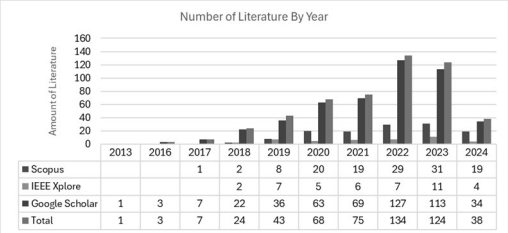
This research is carried out using systematic literature review. It is necessary to compare the existing types of systematic reviews, including meta-analysis, mixed methods study and scope review. Meta-analysis studies are studies using similar datasets and within a smaller study scope. Determining the dataset and scope of the study is based on a set of predetermined criteria. Some examples of criteria that can be used include the language used in the research, the way the research is carried out, sample selection criteria and so on. To obtain greater similarities and conclusions, authors can use statistical methods to analyze existing studies. Significant literature reviews and data collection by authors can be carried out as studies using mixed methods. To analyze the data, the authors used quantitative and qualitative methodology. A scoping review study is also carry out to determine the scope of the review by determining how much the avalibility of information related to the research questions and research topic. Sometimes this study are used as preparation for a larger systematic review [11].

The study carried out in this paper aims to look at research developments and potential future research opportunities based on research samples that have been selected based on certain criteria. Therefore, the systematic review and meta-analysis method will be used. The format for preparing this research report refers to the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) standard. To increase the validity and standardization of carrying out systematic reviews and meta-analysis, researchers created PRISMA. And these standards and recommendations are regularly updated [11].

To determine the range of years that will be used as material for conducting a literature review, an analysis of the range of years of research related to blockchain and data governance is carried out. Searches were carried out using the words blockchain and data governance in the title, abstract or keywords. Search results can be seen in Figure 1. Then an analysis of the impact of these articles was carried out by counting the number of citations.

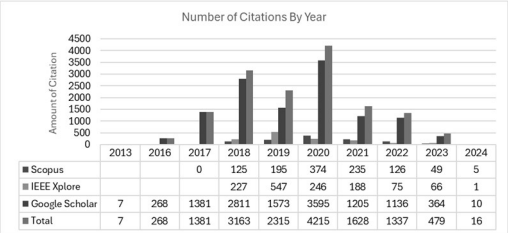
It can be seen in Figure 1 and Figure 2 that the research related to blockchain and data governance was started to increase in 2017. It can also be seen in Figure 1 that the initial articles regarding blockchain and data governance appeared in 2013 with a total of 7 citations. Based on the number of articles, citations and research recency in Figure 2 and Figure 3, it is decided that the research articles that would be used

to carry out the systematic review are articles published starting in 2017 until 2024.



	2013	2016	2017	2018	2019	2020	2021	2022	2023	2024
Scopus			1	2	8	20	19	29	31	19
IEEE Xplore				2	7	5	6	7	11	4
Google Scholar	1	3	7	22	36	63	69	127	113	34
Total	1	3	7	24	43	68	75	134	124	38

Figure 2: Number of Literature By Year



	2013	2016	2017	2018	2019	2020	2021	2022	2023	2024
Scopus			0	125	195	374	235	126	49	5
IEEE Xplore				227	547	246	188	75	66	1
Google Scholar	7	268	1381	2811	1573	3595	1205	1136	364	10
Total	7	268	1381	3163	2315	4215	1628	1337	479	16

Figure 3: Number of Citation By Year

The literature review was carried out using PRISMA [12]. PRISMA also is a method for including or excluding articles [13]. PRISMA provides a structured and systematic flow, making it easier to conduct a literature review [14]. Thus, it can be used to investigate a study [15].

The literature review is carried out through the following stages:

1. Step 1: Search articles on Scopus, IEEExplore, Science Direct and Google Scholar
   The search for these articles was carried out based on the suitability of the contents of the articles with the specified keywords, inclusion criteria and exclution criteria.
2. Step 2: Analyzing research trends
   Analyze the increase in research and development patterns of related research topics
3. Step 3: Mapping research results with predetermined research questions.
   Carrying out mapping analysis of research results contained in articles with predetermined research questions.
4. Step 4: Synthesizing and concluding
   Analyzing blockchain-based data governance patterns and utilization of blockchain in data governance. As well as concluding potential future research.

2.1. Search articles by keywords

At this stage, articles are searched based on certain keywords. Keywords used in searches include:

"blockchain"
"blockchain based", "blockchain-based"
"data governance"
"data governance framework"

The number of research articles obtained based on these keywords can be seen in the Table 1.

Table 1: Number of research articles obtained based on keywords

Keywords/Keyword combinations	Database	Number of results
blockchain AND "data governance"	Scopus	76
blockchain AND "data governance"	IEEExplore	84
"blockchain" "data governance"	Science Direct	151
"blockchain" or "blockchain-based" or "blockchain based" and ("data governance" or "data governance framework")	Google Scholar	138

From the literature search process based on these keywords, a total of 449 research articles were obtained.

2.2. Inclusion and Exclusion Criteria

The process continues referring to the PRISMA protocol. The process of analyzing the 449 articles obtained in the first step was continued by applying certain criteria. There were 44 duplicate articles that were not included. Apart from that, to increase the relevance of these articles, filtering is carried out based on other additional criteria. Other additional criteria include articles published in English and within the time period between 2017 - 2024. Apart from that, there are several additional criteria, including the type of document selected, namely the type of article document whose publication process has reached the final step, as well as articles that openly accessible which are Open Access & Open Archive and also the title contains the word framework. There were 405 articles that met the specified criteria. This flow is depicted in Figure 3 in the Identification step.

After that, a screening step was carried out by reading the titles and abstracts of the articles that had been identified. This is done to determine the suitability of the article with the research topic of this paper. The results of this step can be seen in Figure 3, where 41 suitable articles were obtained.

The next step is to carry out an analysis of the eligibility level of the existing articles by reading all 41 articles. The results of this step can be seen in Figure 3. Eligibility step, where 34 articles are considered eligible and 7 articles cannot be accessed.

The final step in PRISMA is to determine which articles will be included as material to contribute to this paper. This step was carried out by looking at the relationship between 34 articles and the research question. The results of this step can be seen in Figure 3 which 15 articles included in the analysis material. The list of 15 articles can be seen in the table 6.

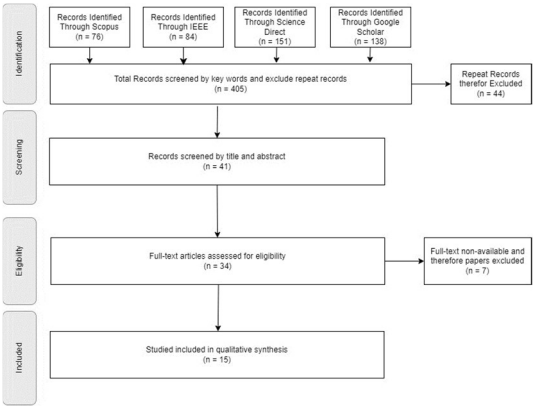


Figure 4. The PRISMA diagram depicts the study selection and identification process

2.3. Research Questions

This section discuss about research questions that will be answered through a systematic literature review process. These research questions were chosen by adopting 5W1H (what, why, where, when, who and how). These research questions can be seen in the Table 2.

Table 2: Research Questions and motivations

No	Research Questions	Motivation
RQ1	What aspects are the focus of blockchain utilization for data governance? (What)	To observe how primary studies define the concept of data governance through blockchain.
RQ2	Why is data governance adopted through blockchain? The background and motives for the utilization of blockchain for data governance. (Why)	To understand the forces of data governance through blockchain.
RQ3	Where is blockchain enforced data governance? The components that blockchain play in implementing data governance. (Where)	To distinguish the key blockchain objects for data governance.
RQ4	When is data governance with blockchain applied? (When)	To understand where blockchain-based data governance fits in.
RQ5	Who is involved in blockchain for data governance? (Who)	To identify the different roles, and their authorities, capabilities and responsibilities in blockchain for data governance.
RQ6	How is data governance through blockchain designed? (How)	To explore actionable mechanisms for implementing blockchain-based data governance.

3. RESULT AND DISCUSSION

3.1. Research Trends

Research trend analysis of 405 articles obtained from the first step of PRISMA can be seen at Figure 5. From the 405 articles, it can be seen that research related to blockchain and data governance has increased. Therefore, it can be concluded that blockchain and data governance are two areas that are increasingly relevant and needed.

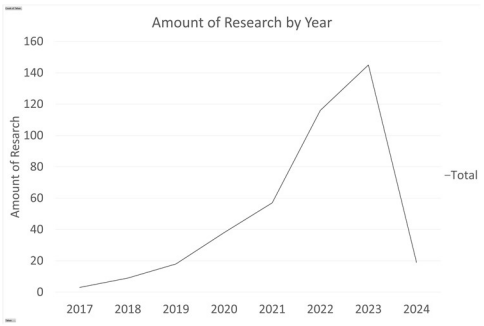


Figure 5: Research trends related to blockchain and data governance from 2017-2024

Figure 6 provides an overview of the trend from 34 articles. In 2020 research regarding the utilization of blockchain in data governance was touch on aspects of security and privacy. Then in the following years research developed to look at the relationship between blockchain and data governance with aspects of trust, data quality and cybersecurity. It can conclude that there are opportunity to conduct research and development related to the utilization of blockchain-based data governance in improving trust, data quality and cybersecurity. The requirement to implement regulations related to data such as the GDPR in the EU, CPRA in the US and PDP in Indonesia can also encourage the use of blockchain in enforcing data security, data trust and data governance. Blockchain plays a role in maintaining distributed and decentralized transaction records with security and is based on the agreement of relevant stakeholders.

In the future, organizations are required to be able to collaborate globally, where the stakeholders involved will become wider. Furthermore, organizations are required to be able to make decisions quickly and accurately. So the opportunity for organizations to become a Decentralized Autonomous Organization (DAO) is even greater. Therefore, managing data security, data privacy and data trust between related parties is increasingly critical. This is where blockchain can play a role.

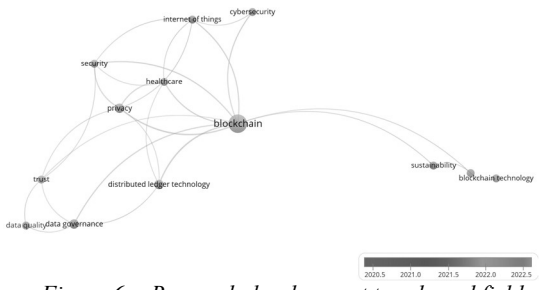


Figure 6: Research development trends and field interrelationships

3.2. Results and Discussion

This section discusses the analysis results from 15 articles research obtained from the fourth step of PRISMA. These articles are used to answer the research questions. Table 3 explains that the use of blockchain for data management has developed. Not only in fields related to crypto currency, but it has been used in other fields such as health, academic management, population management, environmental management and so on.

Table 3: Main information of articles on blockchain and data governance framework

Description	Result
Timespan	2017-2024
Domain used for research implementation	Big data, Supply chain, Scientific Research Publishing, Healthcare, Digital Economy, Refugees Document Authentication, IoT, Smart City Development, Waste Management, Voting Systems
Technology Used	Truffle, Go Language, Petri Net Markup Language, React. js, Metamask, Node. Js, Express, Ethereum, Sidechain, Edge Computing, Go-Ethereum, Medical IoT Device, Off-Chain Distributed Storage, Smartcontract, Interplenatry File System (IPFS) Protocol, Virtual Machine (Ubuntu OS), Proof of Authority (PoA) consensus protocol (Clique), Solidity, Proxy Re-Encryption Mechanism (PRE), Python (NuCypher pyUmbral), Ganache, Springboot, MySQL, Multichain, JSON, Axios, Geth, Infura, Ropsten

3.2.1. RQ1: What is data governance through blockchain?

This research question (RQ1) aims to obtain an overview of the concept of data governance carried out through the application of blockchain. To answer this research question, several literatures provide explanations regarding the concept of using

blockchain related to data governance. Li, Z., Liang, F., & Hu, H., [17] stated that in terms of data governance, blockchain is used as a means to implement a consensus mechanism so that evaluation of the value of data can be seen from various points of view, and carried out in balance. Apart from that, blockchain also builds standards for sharing and transforming data between different companies/organizations. Blockchain is also used to develop process modeling mechanisms based on the value of data that are compatible with various types of business interactions both within the company and with parties outside the company. Duzha *et al.,* [27] stated that blockchain also provides organizations with the ability to make changes to organizational structures and create transparent distributed governance with a leaner hierarchy compared to traditional organizational structures. Akkaoui *et al.,* [28] stated that blockchain can be used in sharing health data, because blockchain provides an efficient and secure framework. Ezzine *et al.,* [29] stated that blockchain store hashing of data that pointing to a database containing the reliable data. With blockchain's ability to validate digital identities when storing, accessing and sharing data, than data governance that is more compatible with distributed network schemes can be developed [30].

3.2.2. RQ2: Why is data governance adopted through blockchain?

This research question (RQ2) aims to gain an understanding of the impetus for using blockchain in implementing data governance. Sasikumar *et al*., [16] stated that the blockchain system can create an environment that can exchange data and connect with big data in a transparent, open, safe and reliable manner. Furthermore, Li *et al*., [17] stated that blockchain can collaborate enterprises and ensure the confidentiality of the data they share, as well as increase trust between these enterprises. Blockchain also provides a mechanism for reconciling differences in data values from several participants and provides the ability to reconcile the workflow of the collaborating enterprises. Mackey *et al.,* [18] stated that blockchain has the capability to prevent data falsification and fraud, as well as ensuring that public trust is maintained. Mackey *et al*., [18] stated that blockchain enable data provenance, increased transparency, and enhanced trust within a distributed network and smart contract in blockchain can help in achive agreement and validate transactions among the distributed nodes. Karisma *et al*., [21] stated that blockchain has an edge in data privacy and it provides transparency across multiple chain and prevents data modification with embedded feature of immutability. Akkaoui *et al*., [28] stated that by using merkle tree as the basis of the blockchain, data changes can be prevented. Because data changes will cause changes to other interrelated blocks. Blockchain provides the ability to manage transactions from trustless individuals, without the involvement of a trusted third party (TTP). These transactions are carried out safely, reliably, transparently, pseudoanonymously, transparently and trustworthy. Furthermore, blockchain also has the capability to execute consensus protocols, so that transactions are certified, mined and approved first before the transaction record is added to a block which is linked to the previous block. Ezzine *et al.,* [29] stated that blockchain technology has the ability to remain unchanged and indelible, to keep authorized users responsible for any transaction and to share data with appropriate authentication without third party intervention. Blockchain can bring solution to the privacy and security problems of P2P networks. By enabling private key management, cryptography and self-sovereign identity, as well as the distributed nature of blockchain, it can increase the transfer, sharing and access of healthcare data quickly, transparently and standardized [30].

3.2.3. RQ3: Where is blockchain enforced data governance?

This research question (RQ3) aims to distinguish the key blockchain objects for data governance. Data validation using blockchain is carried out using a hashing algorithm. The hasing algorithm summarizes information about data elements, hash values, data sources and operations performed, which then becomes information about the collaboration required [17]. Furthermore, in blockchain, the value of the data is evaluated through an auction mechanism. Akkaoui *et al*., [28] Blockchain is used to ensure privacy and security in data sharing. And if someone tries to change a transaction from the past, the hash value of the block that contains the transaction must be calculated again and all hash values for the blocks that come afterwards must be calculated again too. This can't be done, unless more than half of the nodes in a network are infected [29]. Blockchain and off-chain blockchain system components were designed to have a verified history of all entities that processed or accessed their protected information in a compliant fashion and to enable data owner control and access to their data from off-chain data sources and also [30].

3.2.4. RQ4: When is data governance with blockchain applied?

This research question (RQ4) aims to understand where blockchain-based data governance fits in. Data governance based on blockchain can be used when there is a requirement for implementing collaborative production that involving various data elements need, when privacy protection need to be adopted. Besides that blockchain is also used in industries that require the involvement of various parties in the work chain. In this case, the blockchain verifies the authentication and consistency of data in the database originating from various parties without authority from management [17]. Sasikumar *et al*., [16] using blockchain for data management when secure access to data is needed. For information that is private and sensitive, blockchain can be used to manage strong authorization and authentication procedures to access that information [28]. Blockchain also used when there is a vulnerability against cyberattacks and when there is a need for collaborative and distributed data governance. Ezzine *et al.,* [29] stated that blockchain used when there is a need to improve the quality of data. Data governance applied with blockchain applied when there is a system that relatively complex, has regulatory requirements, and by nature involving multiparty [30].

3.2.5. RQ5: Who is involved in blockchain for data governance?

This research question (RQ5) aims to identify the different roles, and their authorities, capabilities and responsibilities in blockchain-based data governance. From a review of several literatures, it can be concluded that the various parties involved in using blockchain for data management consist of those who own the data, those who use the data and those who monitor the data [20, 22-23]. Akkaoui *et al*., [28] proposed blockchain based framework which collaborates end users as data generators and data requestors. Ezzine *et al*., [29] stated that blockchain based data governance framework involving with source of information. Miyachi *et al.*, [30] proposed framework that collaborates patient as data owners, administrators and intermediaries.

3.2.6. RQ6: How is data governance through blockchain designed?

This research question (RQ6) aims to explore actionable mechanisms for implementing blockchain-based data governance. Sasikumar *et al*., [16] stated that each block in the blockchain store a timestamp, collective signature, and the hash produced by the data. To verify the block, it is done using a consensus algorithm. The requirement to include various parties can lead to varying perspectives regarding the value of a data element, smart contracts can be used to coordinate these different perspectives. So that data interaction at the enterprise level can be achieved. The data evaluation mechanism on the blockchain uses an auction model. Evaluation of the key attributes of the data owner is carried out by assessing the price per unit of data, how much data is represented and sub-attributes of data quality. Blockchain are set up so that the collected data can be defined, configured, analyzed, verified and broadcast. Furthermore, hashing algorithms are used to ensure the authenticity of each data message collected from the source and measure the integrity of data elements [17]. On the blockchain, we can also set rules and permissions for health data access and endorsements to data requestors using smart contracts. They also using the Proof of Authority (PoA) consensus mechanism for the validation of new blocks and a blockchain database for storing the hashes of the EMRs generated using keccak-256 as well as URL hash pointers [28]. Ezzine *et al.*, [29] using blockchain for containing hashes of data that pointing to a database containing the reliable data. So there will be a trusted blockchain that refers to the data in the database. Smart contracts on the blockchain technology governed the the timestamped tracking of any existing data securely. The transfer, access control, and also authorization of data controlled through smart contracts [30].

3.3. Discussion

Based on the answers to the research questions, it can be concluded that data governance via blockchain is carried out in processes related to data value evaluation, data standardization and transformation, data transparency, data security, data sharing and data storage. The use of blockchain in implementing data governance is carried out because blockchain has the capability to create a transparent, open, safe and reliable environment. Apart from that, blockchain is able to collaborate with various parties while maintaining data confidentiality. Blockchain has the capability to reconcile differences in data values thereby preventing data falsification and fraud. The use of blockchain in implementing data governance is carried out because blockchain has the capability to create a transparent, open, safe and reliable environment. Apart from that, blockchain is

able to collaborate with various parties while maintaining data confidentiality. Blockchain has the capability to reconcile differences in data values thereby preventing data falsification and fraud. So that public trust can be maintained.

Blockchain is used in the process of evaluating data value as well as in the process of recording and maintaining the history of data changes. The use of blockchain in data governance is carried out in an environment that requires collaborative work processes, involves diverse data and requires aspects of data privacy. The parties involved in implementing blockchain-based data governance include data owners, data users, parties who have the obligation to monitor data management process and other intermediary parties who carry out data processing or data transmission. Implementing data governance using blockchain is carried out by recording timestamps and data signatures. Apart from that, data management is carried out by utilizing data hashing mechanisms, the use of algorithm in smart contracts and consensus mechanisms.

4. CONCLUSIONS

Based on the articles have been selected, it can be concluded that the development of researches are focusing on the utilization of blockchain to answer aspects of data governance such as data evaluation process, establishing standards for sharing and transforming data, establishing interaction models and data transparency. Furthermore, blockchain also answering data protection and privacy aspect. Blockchain also plays a role in data validation by utilization hashing algorithm. Data validation is carried out on data elements, including data sources and information about operations carried out. Analysis on frameworks proposed from the 15 articles conclude that the parties involved in the blockchain are data providers, data consumers and those who monitor the data. Data governance is carried out via blockchain by implementing smart contracts to coordinate different parties, hashing algorithm to ensure the authenticity, consensus mechanism to make sure the finality of the block. In general, it can be concluded that to utilize blockchain in implementing data governance it needs to be determined the smart contract algorithm for ensuring the validity, security and privacy of data, data management and transformation processes and collaboration patterns for the parties involved.

The major components that form blockchain-based data governance, as in Figure 7, are users, access and integration protocol, and data storage. Users as data owner, data producer or data source and user as data consumers can be in the form of human, or surrounding applications with databases. Access and integration protocols are used for validating the data, validating data access/authenticating. It can be in the form of integration layer using API or smart contract in the blockchain platform. Data storage can be in the form of database in data center or in blockchain platform.

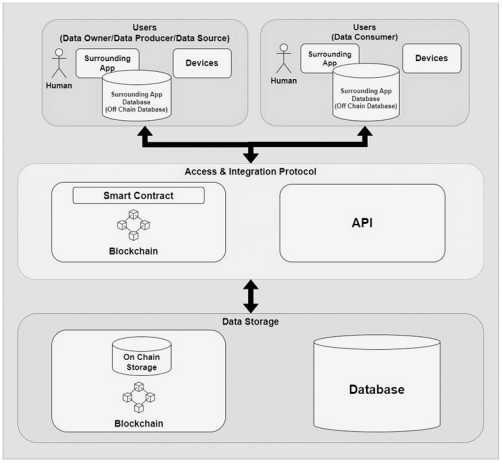


Figure 7: Components of Blockchain-based Data Governance

Based on the literature review, there are several future researches that can be carried out, including those related to development of a consensus mechanism that uses energy efficiently, research on statistical models to determine the size of modules for data sharing so that they can meet data sharing requirements and match optimal block size and upload times. Furthermore, research related to implementation of other variations of data privacy technique such as k-anonymity, t-closeness and i-diversity. Research on the use of artificial intelligence and machine learning on blockchain. Research on algorithm for better real-time decision, by training the system to proactively predict intelligent insights. Other research related to environment and economy such as natural environment design that promote environmental sustainability, token economy and data decentralization movements with economic empowerment. Other important research for the future is related to interoperability of the blockchain and the security of smart contracts. Research related to data governance frameworks such as implementation of Master Data Management (MDM) policies with blockchain, research in which Decentralized Autonomous Organizations (DAOs)

are a suitable governance framework for data, data governance across a supply chain related to preponderance of IoT devices.

This research work focus on the role of blockchain and the components of blockchain-based framework related to data governance but not the algorithms used to ensure data trust as an important part of ensuring good data governance.

In future the researchers can work on the improvement of data trust algorithm in smart contracts and blockchain-based framework that pay attention to the data governance and IT governance framework. So it can support in ensuring better data governance implementation and the maturity level of IT governance in the organization.

REFERENCES

[1] Statista.com. (2023), https://www.statista.com/statistics/871513/worldwide-data-created/

[2] Truechart.com. (2023), https://www.truechart.com/benefits-of-data-collaboration-for-your-team/

[3] Dama International, (2017), "DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition)". USA:Technics Publications, LLC.

[4] Data Governance Institute. "Definitions of Data Governance." The Data Governance Institute, August 10, 2023. https://datagovernance.com/the-data-governance-basics/definitions-of-data-governance/.

[5] Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf

[6] Merriam-Webster.com. (2024). https://www.merriam-webster.com

[7] Zheng, Zibin, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang. 2017. "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends." Proceedings - 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017, no. October: 557–64. https://doi.org/10.1109/BigDataCongress.2017.85.

[8] Tripathi, Gautami, Mohd Abdul Ahad, and Gabriella Casalino. 2023. "A Comprehensive Review of Blockchain Technology: Underlying Principles and Historical Background with Future Challenges." Decision Analytics Journal 9 (March): 100344. https://doi.org/10.1016/j.dajour.2023.100344.

[9] Ghosh, Paramita. "The Fundamentals of a Data-Driven Approach." Dataversity. February 28, 2023. https://www.dataversity.net/the-fundamentals-of-a-data-driven-approach/.

[10] Hinkle, O. "The Evolution of Data Governance". May 18, 2020. https://www.dataversity.net/the-evolution-of-data-governance/

[11] Fulton Library. "Research Methodologies," Accessed May 5, 2024. https://uvu.libguides.com/methods/systematic-review

[12] L. Shamseer, D. Moher, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation," BMJ, vol. 349, no. jan02 1, pp. g7647–g7647, jan 2015, doi: http://dx.doi.org/10.1136/bmj.g7647.

[13] Kanga Id´e Soumaila, Albachir Seydou Niandou, Mustapha Naimi, Chikhaoui Mohamed, Keith Schimmel, Stephanie Luster-Teasley, and Naeem Nizar Sheick, "A Systematic Review and Meta-Analysis of Water Quality Indices," Journal ofAgricultural Science and Technology B, vol. 9, no. 1, pp. 1–14, jan 2019, doi: http://dx.doi.org/10.17265/2161-6264/2019.01.001.

[14] L. A. Stewart, M. Clarke, M. Rovers, R. D. Riley, M. Simmonds, G. Stewart, and J. F. Tierney, "Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data," JAMA, vol. 313, no. 16, pp. 1657–1665, apr 2015, doi: http://dx.doi.org/10.1001/jama.2015.3656.

[15] S. Lee, Y. Lee, S. Choe, and W. Lee, "Adverse Sexual Effects of Treatment with Finasteride or Dutasteride for Male Androgenetic Alopecia: A Systematic Review and Meta-analysis," Acta Dermato Venereologica, 2018, doi: http://dx.doi.org/10.2340/00015555-3035.

[16] Sasikumar, A., Logesh Ravi, Ketan Kotecha, Ajith Abraham, Malathi Devarajan, and Subramaniyaswamy Vairavasundaram. (2023). "A Secure Big Data Storage Framework Based on Blockchain Consensus Mechanism With Flexible Finality." IEEE Access 11 (May): 56712–25.

https://doi.org/10.1109/ACCESS.2023.3282322

[17] Li, Zhi, Fuhe Liang, and Henan Hu. (2023). "Blockchain-Based and Value-Driven Enterprise Data Governance: A Collaborative Framework." Sustainability (Switzerland) 15 (11). https://doi.org/10.3390/su15118578

[18] Mackey, T. K., Shah, N., Miyachi, K., Short, J., & Clauson, K. (2019). A Framework Proposal for Blockchain-Based Scientific Publishing Using Shared Governance. Frontiers in Blockchain, 2(November), 1–12. https://doi.org/10.3389/fbloc.2019.00019

[19] Malik, Varun, Ruchi Mittal, Dinesh Mavaluru, Bayapa Reddy Narapureddy, S. B. Goyal, R. John Martin, Karthik Srinivasan, and Amit Mittal. 2023. "Building a Secure Platform for Digital Governance Interoperability and Data Exchange Using Blockchain and Deep Learning-Based Frameworks." IEEE Access 11 (June): 70110–31. https://doi.org/10.1109/ACCESS.2023.3293529

[20] Mackey, T. K., Calac, A. J., Chenna Keshava, B. S., Yracheta, J., Tsosie, K. S., & Fox, K. (2022). Establishing a blockchain-enabled Indigenous data sovereignty framework for genomic data. Cell, 185(15), 2626–2631. https://doi.org/https://doi.org/10.1016/j.cell.2022.06.030

[21] Karisma, K., & Tehrani, P. M. (2022). Data protection governance framework: A silver bullet for blockchain-enabled applications. Procedia Computer Science, 218(November), 2480–2493. https://doi.org/10.1016/j.procs.2023.01.223

[22] Garcia, R Dutra, G Sankar Ramachandran, and ... 2021. "A Blockchain-Based Data Governance Framework with Privacy Protection and Provenance for e-Prescription." ArXiv e …. https://ui.adsabs.harvard.edu/abs/2021arXiv211213956D/abstract

[23] Vijayvergiya, Rashmi Mandal, and Soumya Sen. 2023. "A Framework to Maintain Child Immunization Records in Secure Ethereum Blockchain-Enabled Platform" 2 (1): 7–20.

[24] Nabben, K. (2022). Decentralised Autonomous Organisations (DAOs) as Data Trusts: A General-purpose Data Governance Framework for Decentralised Data Ownership, Storage, and Utilisation. SSRN Electronic Journal, 1–14. https://doi.org/10.2139/ssrn.4009205

[25] Liu, Y., Lu, Q., Yu, G., Paik, H. Y., & Zhu, L. (2022). Defining blockchain governance principles: A comprehensive framework. Information Systems, 109(August 2015). https://doi.org/10.1016/j.is.2022.102090

[26] Shrestha, A. K. (2021). Designing incentives enabled decentralized user data sharing framework. Harvest. https://harvest.usask.ca/handle/10388/13739

[27] Duzha, A., Alexakis, E., Kyriazis, D., Sahi, L. F., & Kandi, M. A. (2023). From Data Governance by design to Data Governance as a Service: A transformative human-centric data governance framework. ACM International Conference Proceeding Series, 10–20. https://doi.org/10.1145/3616131.3616145

[28] Akkaoui, Raifa, Xiaojun Hei, and Wenqing Cheng. 2020. "EdgeMediChain: A Hybrid Edge Blockchain-Based Framework for Health Data Exchange." IEEE Access 8: 113467–86. https://doi.org/10.1109/ACCESS.2020.3003575.

[29] Ezzine, I, and L Benhlima. 2021. "Technology against COVID-19 A Blockchain-Based Framework for Data Quality." 2020 6th IEEE Congress on Information …. https://ieeexplore.ieee.org/abstract/document/9357200/.

[30] Miyachi, Ken, and Tim K. Mackey. 2021. "HOCBS: A Privacy-Preserving Blockchain Framework for Healthcare Data Leveraging an on-Chain and off-Chain System Design." Information Processing and Management 58 (3): 102535. https://doi.org/10.1016/j.ipm.2021.102535.

Table 4: Selected articles

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
1.	Sasikumar *et al.*	2023	A Secure Big Data Storage Framework Based on Blockchain Consensus Mechanism With Flexible Finality	Big data storage; blockchain; consensus mechanism; flexible finality; highway protocol	This study introduces a brand-new approach to big data storage security that leverages blockchain technology and applies highway protocol to generate new blocks.	✓	✓	✓	✓	✓	✓	[16]
2.	Li Z.; Liang F.; Hu H.	2023	Blockchain-Based and Value-Driven Enterprise Data Governance: A Collaborative Framework	blockchain; collaborative production; data governance; governance and technology; value-driven	This study focuses on development of a multidimensional and collaborative data governance mechanism between industry chains in a sharing economy.	✓	✓	✓	✓	✓	✓	[17]
3.	Mackey *et al.*	2019	A Framework Proposal for Blockchain-Based Scientific Publishing Using Shared Governance	academic publishing; blockchain; data governance; distributed ledger technology; scientific integrity; scientific research	This study formulates a new blockchain governance framework focused on increasing transparency in scientific publishing.	-	✓	-	-	-	-	[18]
4.	Malik *et al*.	2023	Building a Secure Platform for	Blockchain;data exchang	This study proposes a solid stage	-	-	✓	-	-	-	[1

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
			Digital Governance Interoperability and Data Exchange Using Blockchain and Deep Learning-Based Frameworks	e;deep reinforcement learning; digital governance;interoperability;Internet of Things;voting system; waste management;cybersecurity	framework in view of blockchain and profound figuring out how to increment trust of digital governance interoperability and data exchange.							9]
5.	Mackey *et al*.	2022	Establishing a blockchain-enabled Indigenous data sovereignty framework for genomic data	To develop the early foundations of an Indigenous data sovereignty (IDS) blockchain framework that is community centered and enables Indigenous peoples to engage in distributed sovereign data management.	This study focuses on development of the early foundations of an Indigenous data sovereignty (IDS) blockchain framework that is community centered and enables Indigenous peoples to engage in distributed sovereign data management.	-	✓	-	-	✓	-	[20]
6.	Karisma,	2023	Data protection governance	Blockchain technolo	This study focuses on the development	-	✓	-	-	-	-	[2

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
	Tehrani.		framework: A silver bullet for blockchain-enabled applications	gy; data protection; governance framework Keywords:	of data protection indicators (DPIs) that have a bearing on the legal readiness and maturity levels of countries in developing blockchain-enabled solutions.							1]
7.	Garcia *et al*.	2021	A Blockchain-based Data Governance Framework with Privacy Protection and Provenance for e-Prescription	Data Governance, Decentralized, E-prescription, Privacy, Blockchain, Smart Contracts, Proxy Re-encryption	This study focus on proposing a decentralized data governance framework for the electronic prescription that helps patients store, manage, and share prescription data with other stakeholders, protects patients' privacy, provides support for consent management using proxy re-encryption scheme and smart contracts, supports data provenance, uses privacy-enhancing data management scheme to withhold personally identifiable and sensitive	-	-	✓	-	✓	-	[22]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
					information from third parties and enables the drug regulators to control and monitor the flow of medications to the pharmacies through the accountable blockchain ledger.							
8.	Vijayvergiya, R. M., & Sen, S	2023	A Framework to Maintain Child Immunization Records in Secure Ethereum Blockchain-Enabled Platform	Blockchain, Ethereum, Solidity, tamper-proof record, child immunization	This study focus on proposing a novel way of maintaining data using the Ethereum-enabled Smart Contract mechanism to maintain the immunization record of the child on the blockchain.	-	-	✓	-	✓	-	[23]
9.	Nabben, K	2021	Decentralised Autonomous Organisations (DAOs) as Data Trusts: A general-purpose data governance framework for decentralised data ownership, storage, and utilisation	Decentralized autonomous organization, data trust, data governance, blockchain	This study focuses on investigation of how principles of data trusts might apply to "Decentralised Autonomous Organisations" (DAOs) as a data governance structure for digital data trusts and explore DAO as a data governance infrastructure that is	-	-	✓	-	-	-	[24]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
					analogous to data trusts.							
10.	Liu *et al*.	2022	Defining blockchain governance principles: A comprehensive framework	Blockchain, governance, decision right, accountability, incentive, ecosystem	This study focuses on proposing a blockchain governance framework that categorises the governance structures according to the decentralisation level of governed blockchains	-	-	-	-	-	✓	[25]
11.	Shrestha, A. K.	2022	Designing incentives enabled decentralized user data sharing framework	Blockchain; Blockchain-based system; Data Trust; Incentives; PLS; Privacy; SEM; Security; Smart contract; Technology Acceptance Model; Trust	• The first research objective is to review the most relevant literature in the field of user datasharing platforms that incentivize users for their contribution to discover gaps in the existing systems and methods towards user-controlled privacy-preserving user data sharing approaches.			✓				[26]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
					• Second objective focuses on the development of a standard definition of user data and classify user data to support all kinds of user-profiles data, user documents (usergenerated/created data) and research data. • Third objective focuses on the development of a decentralized user data sharing framework, sents a series of empirical studies and proposes novel blockchains- and smart contracts-based DUDS (Decentralized User Data Sharing) framework							

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
					• Fourth objective focuses on analyzing the preliminary factors that affect acceptance of a prototype system developed with the DUDS framework • Fifth objective focuses on designing a user model that enables examining the role and dimensions of various antecedents of the behavioral intention of users to adopt the data sharing platforms such as the DUDS platform. • Sixth objective focuses on implementation the DUDS framework • Seventh objective focuses on evaluation							

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
					of the user experience model and the DUDS framework • Eighth objective focuses on the measurement of the performance metric in the private blockchain network along with evaluating the user experience model							
12.	Duzha *et al*.	2023	From Data Governance by design to Data Governance as a Service: A transformative human-centric data governance framework	Data Governance, Data Processing, Data Quality, Decentralization, Trust	• First objective of this study focuses on introducing a novel approach for data governance as a service (DGaaS), which provides a framework for organizations that facilitate alignment with their vision, goals and legal requirements, while at the same time increasing	✓	✓	-	-	-	-	[27]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
					awareness among individuals and teams involved in the value chain. • Second objective on this study focuses on introducing the conceptual architectural of an integrated environment designed to provide trustworthy mechanisms for Data Processing using AI and Big Data analytics.							
13.	Akkaoui R.; Hei X.; Cheng W.	2020	EdgeMediChain: A Hybrid Edge Blockchain-Based Framework for Health Data Exchange	Blockchain; data sharing; edge computing; electronic medical records; healthcare; Internet of Things; privacy; security; smart contracts	• This study focuses on the presentation of a secure and efficient data management framework, named "EdgeMediChain", for sharing health data.	✓	✓	✓	✓	✓	✓	[28]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
14.	Ezzine, I, and L Benhlima	2021	Technology against COVID-19 A Blockchain-Based Framework for Data Quality.	Big Data; Blockchain; Covid-19; Data governance; Data Quality	• This study focuses on highlighting some benefits of using BIG Data and Blockchain to deal with this pandemic and some data quality issues that still present challenges to decision making. • The second objective of this study focuses on presenting a general Blockchain-based framework for data governance that aims to ensure a high level of data trust, security, and privacy.	✓	✓	✓	✓	✓	✓	[29]
15	Miyachi K.; Mackey T.K.	2021	hOCBS: A privacy-preserving blockchain framework for healthcare data	Blockchain; Data storage; Distributed ledger technology;	• This study focuses on reviewing the characteristics of different	✓	✓	✓	✓	✓	✓	[30]

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Ref
			leveraging an on-chain and off-chain system design	Health informatics; Healthcare; Privacy; Secure computation	constructs of Off-Chain Blockchain Systems (OCBS). • Second objective of this study focuses on proposing a modular hybrid privacy-preserving framework. • Third objective of this study focuses on conceptualizing, designing, and evaluating a blockchain-based information management system framework that could simultaneously utilize Off-Chain Blockchain Systems (OCBS) system architecture while also preserving performance and security features							

No	Author	Year	Title of Study	Keywords	Research Objectives	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	Ref
					required by existing healthcare data management systems, while also taking into consideration the compliance needs of different forms of healthcare data. • Fourth objective of this study focuses on conceptualizing a privacy-preserving blockchain system using OCBS characteristics and then translate it to three reference models of different types of healthcare data for assessment of real-world data application.							