Singaporean Journal of Scientific Research(SJSR)
An International Journal (AIJ)
Vol.17.No.1 2025,Pp.41-52
ISSN: 1205-2421
available at :www.sjsronline.com
Paper Received : 04-11-2024 Paper Accepted: 06-01-2025
Paper Reviewed by: 1.Prof. Cheng Yu 2. Dr.Yarab Baig
Editor : Dr. Chen Du Fensidal

BLOCKCHAIN AND BIG DATA: A HYBRID MODEL FOR ENSURING SECURITY AND PRIVACY IN IOT AND SOCIAL MEDIA– A REVIEW

Nirmalgowri K
Research Scholar, Dept. of Computer Science & Engineering,
VELS Institute of Science, Technology and Advanced Studies(VISTAS),
Chennai, India.

Dr.A.Vidhya
Assistant Professor,
Department of Information Technology,
VELS Institute of Science, Technology and Advanced Studies(VISTAS),
Chennai, India.

Abstract

The exponential growth of social media platforms and the Internet of Things (IoT) has led to vast amounts of data being generated and shared across centralized systems, making them vulnerable to security breaches, privacy violations, and data tampering. Traditional methods of safeguarding data in these environments have proven inadequate in addressing modern security challenges. This review explores the integration of blockchain technology and big data to create a hybrid model that enhances security and privacy in social media and IoT ecosystems. Blockchain's decentralized, immutable, and transparent nature, combined with smart contracts and cryptographic techniques, offers significant improvements in data integrity, access control, and privacy protection. The paper discusses the *benefits of using blockchain to secure user-generated content and real-time data from IoT devices, along with challenges such as scalability, performance trade-offs, and regulatory compliance. Future research directions are proposed to optimize blockchain's potential in social media and IoT environments, focusing on interoperability, scalability, and energy efficiency.*

Keywords: Blockchain, Big Data Security, Data Privacy, Decentralized Systems, Social Media, IoT Ecosystems, Smart Contracts, Data Integrity, Access Control, Scalability, Privacy-Preserving Frameworks

Introduction

The exponential growth of data generated by social media platforms and the Internet of Things (IoT) has created unprecedented opportunities for data analysis, business insights, and real-time decision-making. However, this data explosion has also introduced significant challenges regarding data security, privacy, and integrity. Social media platforms, with millions of users sharing personal information daily, have become prime targets for data breaches, misuse of user data, and privacy violations. Similarly, IoT devices that generate real-time data streams are often vulnerable to cyberattacks, unauthorized access, and data tampering due to their interconnected nature.

Traditional centralized systems, commonly employed by social media platforms and IoT networks, are increasingly being questioned for their vulnerabilities, especially as they rely on single points of failure. These systems expose massive volumes of user data to potential breaches and attacks, making it difficult to guarantee data privacy and integrity. In this context, there is a growing need for decentralized, secure, and tamper-resistant data management solutions that can handle the complexities of modern social media and IoT ecosystems.

Blockchain technology, with its decentralized and immutable architecture, has emerged as a promising solution for addressing these challenges. By distributing data across a network of nodes and utilizing consensus mechanisms, blockchain can ensure that data remains secure, unaltered, and transparent. Smart contracts further enhance security by automating access control and enforcing strict privacy rules. In conjunction with big data analytics, blockchain offers the potential to create a hybrid model that ensures the security and privacy of data in social media and IoT environments.

This paper explores the integration of blockchain technology with big data to form a hybrid model that addresses the key security and privacy challenges faced by social media platforms and IoT networks. It reviews current blockchain applications, the benefits of decentralized data storage, and the role of smart contracts in managing access control. Furthermore, it discusses the challenges of scalability, transaction speed, and regulatory compliance in implementing blockchain within large-scale data ecosystems. The paper also identifies potential areas for future research to enhance blockchain's effectiveness in securing sensitive data in social media and IoT settings.

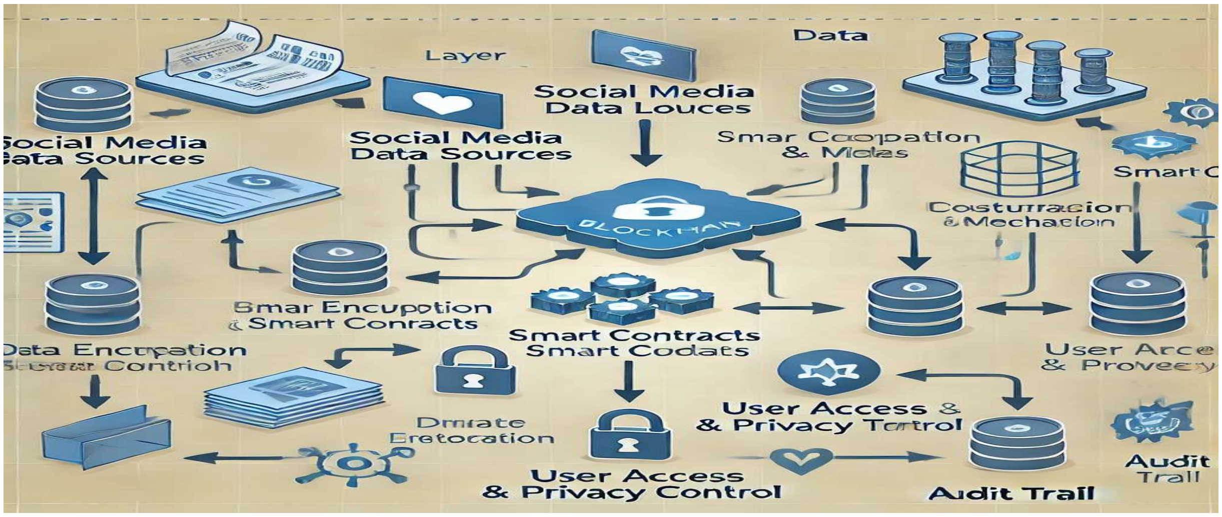

Fig.1 Flow diagram illustrating the integration of blockchain technology with big data for enhanced security and privacy

Here is the flow diagram illustrating blockchain integration with social media for enhanced security and privacy. It visualizes how user data flows through encryption, decentralized storage, and access control mechanisms.

2. Literature Review

Ali et al. (2020) discussed the integration of blockchain and big data for enhancing social media security. They proposed a blockchain-based model to secure social media content, highlighting blockchain's ability to decentralize data storage and prevent tampering with user data. They further demonstrated how smart contracts can be used to control user data access, ensuring that only authorized users can interact with sensitive content.

Zhang et al. (2021) introduced a blockchain-based decentralized approach for securing social media data. The study proposed a framework that utilizes encryption and immutable ledgers to ensure the privacy of user content and activities on social platforms. The researchers argued that blockchain's tamper-proof architecture significantly reduces the risk of data breaches while allowing users greater control over their data.

Rahman et al. (2022) proposed the "Blockchain of Blockchains" framework to ensure IoT data integrity across different networks. They emphasized that blockchain can protect IoT devices from cyberattacks by decentralizing control and distributing data among different nodes, making it difficult for attackers to compromise the system as a whole.

Makhdoom et al. (2020) explored how blockchain can be used for privacy-preserving data sharing in smart cities, which heavily rely on IoT networks. Their research indicated that blockchain not only protects data from tampering but also enhances the transparency of data-sharing practices, ensuring that only authorized entities can access specific IoT data streams. The study concluded that blockchain's decentralized ledger and consensus mechanisms could significantly reduce privacy risks in IoT ecosystems.

Roy et al. (2019) focused on how smart contracts can be used in social media to provide granular access control over user data. Their study showed that smart contracts could automate permission management, allowing users to set conditions for data sharing, thereby enhancing both security and privacy.

Liu et al. (2024) discussed the application of smart contracts in IoT, where they enforce role-based access control to limit unauthorized access to critical device data, ensuring that only permitted users can modify or retrieve sensitive information.

Mitra et al. (2023) explored the impact of blockchain on AI/ML-enabled big data analytics in IoT, discussing the trade-offs between privacy and performance. The study found that while blockchain secures IoT data, it often suffers from slow transaction speeds and high computational overhead as the volume of data increases.

Alhazmi et al. (2022) proposed a big data security framework leveraging blockchain but highlighted the need for improved consensus algorithms to ensure scalability without compromising security.

Younis et al. (2021) discussed how blockchain-based solutions for social media and IoT need to comply with data privacy laws such as the **General Data Protection Regulation (GDPR)**. Their study highlighted the conflict between blockchain's immutability and the "right to be forgotten," a key requirement in GDPR. They proposed solutions such as using off-chain storage for personal data, which could be deleted when needed while still maintaining blockchain's security features for metadata.

Rahman et al. (2022) discussed the importance of finding energy-efficient consensus mechanisms to enhance the scalability of blockchain in IoT environments. They analyzed Proof of Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) as alternatives to Proof of Work (PoW), which is known for its high energy consumption. These newer mechanisms provide faster transaction processing while maintaining a high level of security.

3. Research Gap

Despite the growing interest in applying blockchain to enhance the security and privacy of big data environments such as social media and IoT, several critical gaps remain that need to be addressed for the widespread adoption of these technologies.

3.1. Scalability and Performance Limitations:

One of the most significant challenges identified in the literature is blockchain's scalability. As the size of data generated by social media platforms and IoT devices continues to grow exponentially, blockchain systems face difficulties in handling large volumes of transactions. While consensus mechanisms like Proof of Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) offer some improvements over traditional Proof of Work (PoW) systems, they still struggle to scale efficiently in high-data-volume environments. There is a lack of research on designing blockchain architectures that can handle real-time data processing and high-frequency transactions without sacrificing performance or security, particularly in large-scale IoT and social media ecosystems.

3.2. Privacy vs. Performance Trade-offs:

Although blockchain is lauded for its privacy-enhancing features, such as encryption and Zero-Knowledge Proofs (ZKPs), these techniques introduce significant computational overhead. Existing research has not sufficiently addressed how to minimize the performance degradation caused by privacy-preserving mechanisms, particularly in real-time IoT and social media platforms where low-latency data access is crucial. More research is needed to optimize these privacy mechanisms to ensure they can be deployed without negatively impacting the system’s overall performance.

3.3. Integration with Existing Big Data Systems:

Blockchain is typically deployed in isolation from existing big data frameworks like Hadoop or Apache Spark. The literature lacks extensive studies on how blockchain can be fully integrated into these platforms to provide seamless security and privacy features without affecting the efficiency of big data analytics. Current research is limited in terms of understanding how blockchain can be used to secure large-scale data analytics processes without introducing significant bottlenecks. There is a gap in exploring hybrid models that combine blockchain’s security features with the computational efficiency of established big data systems.

3.4. Regulatory and Compliance Challenges:

Another key research gap involves the conflict between blockchain’s immutability and data privacy regulations, such as the General Data Protection Regulation (GDPR), which mandates the right to delete personal data. Although some studies have proposed off-chain storage or hybrid models to address this issue, these solutions are not yet well-explored or validated in real-world social media and IoT use cases. More research is needed to develop blockchain models that can balance immutability with compliance requirements for data deletion and modification, particularly in highly regulated industries and data-sensitive environments.

3.5. Energy Efficiency:

Blockchain’s energy consumption, particularly in Proof of Work consensus algorithms, remains a significant barrier to its scalability and sustainability. Although alternative consensus mechanisms like PoS and PBFT are more energy-efficient, there is still limited research on optimizing blockchain’s energy consumption for IoT and social media applications. Given the volume and velocity of data generated by IoT devices and social media platforms, more studies are required to design blockchain systems that can handle these data flows without incurring excessive energy costs, making them viable for large-scale, real-time applications.

3.6. Real-Time Data Processing:

In IoT applications, real-time data processing is critical for decision-making and monitoring. However, blockchain's validation and consensus processes introduce delays that make it challenging to support real-time processing. While there are proposals to use sidechains or off-chain processing to accelerate transaction speeds, there is insufficient research on how these solutions can be practically implemented without compromising the security and privacy benefits of blockchain. The development of real-time blockchain solutions that balance security, speed, and privacy in high-frequency data environments remains an underexplored area.

3.7. Limited Research on Use Case-Specific Solutions:

Most blockchain research is generalized and does not focus on the specific needs of different industries or applications. In social media, the primary concern is securing user-generated content and interactions, while in IoT, the focus is on protecting real-time sensor data and device communications. There is a lack of industry-specific blockchain frameworks tailored to address the unique security and privacy challenges in social media and IoT. Research should focus on developing specialized blockchain solutions that cater to the specific characteristics and data flows of these environments.

4. Scope of the Research

The scope of this research focuses on investigating the integration of blockchain technology with big data to address security and privacy challenges in **social media** and **Internet of Things (IoT)** environments. Given the significant data volumes and security risks inherent in these domains, blockchain's decentralized, immutable, and cryptographic capabilities offer promising solutions for protecting user data, ensuring privacy, and maintaining data integrity. The key areas of this research are outlined below:

4.1 Blockchain Integration with Social Media: The study examines how blockchain can enhance the security of user-generated content on social media platforms by decentralizing data storage, preventing tampering, and improving privacy protection. The research will explore the application of **smart contracts** for automating access control, allowing users to manage permissions and control how their data is shared. Additionally, the study will investigate the use of blockchain to mitigate data breaches and unauthorized access in social media ecosystems.

4.2 Blockchain in IoT Security: The research will analyze how blockchain can be integrated with IoT devices to protect real-time data streams, device communication, and sensor-generated data. It will explore blockchain's ability to provide **end-to-end encryption**, ensure secure device authentication, and safeguard the integrity of IoT data through immutable ledgers. The study will also focus on how blockchain can support the growing demand for data integrity, transparency, and tamper-proof systems in IoT networks, particularly in sectors like smart cities and industrial IoT.

4.3 Privacy Protection: The research will evaluate how blockchain can enhance privacy protection in both social media and IoT ecosystems by implementing cryptographic methods such as **Zero-Knowledge Proofs (ZKPs)** and **Homomorphic Encryption**. It will examine how these techniques can safeguard sensitive user information and data streams, allowing only authorized access to specific data without compromising the overall performance of the system.

4.4 Scalability and Performance: The study will assess the **scalability** challenges blockchain faces in handling large volumes of data and high-frequency transactions, particularly in social media and IoT environments where real-time data processing is critical. The research will focus on evaluating different **consensus mechanisms** (e.g., Proof of Stake, Practical Byzantine Fault Tolerance) to address the trade-off between security and transaction speed, providing insights into how blockchain can scale efficiently in data-intensive applications.

4.5 Data Integrity and Access Control: The research will investigate how blockchain can ensure the **integrity** of data generated by social media users and IoT devices by using decentralized consensus mechanisms to verify and validate transactions. The study will also explore how **smart contracts** can be utilized to automate **role-based access control (RBAC)**, allowing for more precise and granular management of data access and permissions across multiple users and devices.

4.6 Energy Efficiency and Resource Optimization: As blockchain's energy consumption remains a challenge, particularly in **Proof of Work** consensus algorithms, this research will examine more energy-efficient alternatives such as **Proof of Stake** and **Delegated Proof of Stake (DPoS)** to ensure blockchain's viability for large-scale deployment in social media and IoT environments. The study will explore how blockchain can be optimized for energy efficiency and resource management without compromising security or privacy.

4.7 Regulatory Compliance: The research will assess how blockchain can comply with data privacy regulations such as the **General Data Protection Regulation (GDPR)** and other legal requirements for protecting user data in social media and IoT networks. It will explore potential solutions to reconcile blockchain's immutability with regulatory demands, particularly the "right to be forgotten" and data deletion requirements.

4.8 Industry-Specific Use Cases: The study will focus on **industry-specific applications** of blockchain, providing case studies from social media and IoT. In social media, the research will investigate how blockchain can protect user interactions, content, and personal data, while in IoT, the focus will be on securing sensor data, device communication, and real-time analytics.

5. Objectives of the Research

- To investigate how blockchain technology can enhance data security and integrity in social media and IoT ecosystems.
- To evaluate the role of smart contracts in automating access control and privacy management.
- To address scalability challenges in blockchain for large-scale data environments like social media and IoT.
- To assess the effectiveness of blockchain in providing privacy-preserving solutions for social media and IoT data.

- To explore the integration of blockchain with existing big data frameworks to optimize performance and security
- To identify energy-efficient solutions for blockchain in high-volume data environments
- To evaluate the regulatory and compliance aspects of blockchain in social media and IoT.
- To propose a framework for blockchain's application in securing social media interactions and IoT data.

6. Experimental Setup

6.1 Data Environment

The experiment will simulate two types of environments. The first environment will represent a social media platform, with user-generated content such as posts, messages, images, and videos. It will also track interactions like likes, comments, and shares, and include user metadata to evaluate the privacy controls in place. The second environment will represent IoT ecosystems, with real-time data streams generated from IoT devices, such as smart city sensors or industrial equipment, producing continuous data flow. The experiment will test how blockchain handles the high-frequency data updates typical in IoT.

6.2 Blockchain Platforms

Three blockchain platforms will be employed for testing: Ethereum, Hyperledger Fabric, and Corda. These platforms were selected due to their suitability for different blockchain use cases—Ethereum for its robust smart contract capabilities, Hyperledger Fabric for its permissioned blockchain and enhanced access control, and Corda for its strong focus on privacy and interoperability. Each platform will be evaluated for its performance in securing social media and IoT data while ensuring privacy and scalability.

6.3 Blockchain Layer

The blockchain layer will include several key components. Smart contracts will be implemented to manage user permissions, content sharing, and data access for both social media and IoT environments. Data will be encrypted using cryptographic techniques to ensure privacy and integrity before storage on the blockchain. Different consensus mechanisms, such as Proof of Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT), will be tested to compare their performance in validating transactions and maintaining blockchain security.

6.4 Data Storage

Data from social media platforms and IoT devices will be stored on a decentralized ledger, ensuring that all actions and data updates are immutable and verifiable. The experiment will also integrate decentralized cloud storage solutions like IPFS or Storj to store large multimedia files (e.g., images, videos), while blockchain will store metadata and enforce access control. This allows for efficient data management while ensuring security.

6.5 Data Access & Privacy

To control access, Role-Based Access Control (RBAC) will be enforced through smart contracts, allowing access based on user roles (e.g., social media friends, followers; IoT administrators, device managers). Additionally, privacy-preserving techniques such as Zero-Knowledge Proofs (ZKPs) and Homomorphic Encryption will be used to ensure that sensitive data remains private and only accessible to authorized users, maintaining compliance with data privacy regulations.

6.6 Scalability Testing

Scalability will be tested by increasing the volume of data in both environments, from small datasets (megabytes) to large ones (terabytes). The network size will also be varied, with simulations ranging from networks with 10 nodes to larger networks with 500 or more nodes. The experiment will track how the blockchain platforms handle this increasing load and measure transaction speed, focusing on how efficiently new data can be validated and appended to the blockchain.

6.7 Performance Metrics

The key performance metrics measured in this experiment will include data integrity, assessing blockchain's ability to prevent tampering; privacy, evaluating the effectiveness of encryption and smart contracts; and transaction speed, tracking the time required to process user interactions or data updates. Additionally, scalability will be evaluated based on how well each platform performs with increasing data volume and node count. Finally, cost and energy efficiency will be tracked by measuring the energy consumption and resource usage of each blockchain platform, comparing their performance in terms of overall efficiency.

7. Hypotheses

The following hypotheses will be tested:

H_1: Blockchain technology significantly improves data integrity and security in social media and IoT environments compared to traditional centralized systems.

Blockchain's decentralized nature, with its immutable ledger, is expected to provide superior data integrity and security when compared to traditional centralized systems used in social media platforms and IoT networks. In a centralized setup, a single point of failure can lead to data tampering, breaches, and unauthorized modifications. In contrast, blockchain's distributed consensus mechanisms ensure that any changes or transactions are verified and recorded immutably, thereby enhancing the overall security and integrity of user-generated content and IoT data streams.

H_2: Blockchain-based smart contracts enhance access control and privacy protection for social media users and IoT data streams.

Smart contracts in blockchain can enforce Role-Based Access Control (RBAC), allowing for more precise control over user permissions and data access in both social media and IoT environments. It is hypothesized that by utilizing smart contracts, data owners can define strict rules regarding who can access or modify their data, thus ensuring better privacy protection. This approach offers an improvement over conventional access control mechanisms, which often rely on less flexible, centralized systems that are prone to breaches and unauthorized access.

H_3: The scalability of blockchain systems is negatively impacted by increased data volumes and network size, resulting in lower transaction speeds and higher resource consumption.

Although blockchain provides strong security features, its performance may degrade as the size of the network and volume of data increase, especially in data-intensive environments like IoT and social media. It is hypothesized that larger data volumes and more network nodes will lead to a decrease in transaction throughput and an increase in latency. Additionally, the increased computational demand for validating transactions across multiple nodes may result in higher energy consumption and slower performance.

H_4: Privacy-preserving techniques such as encryption and Zero-Knowledge Proofs (ZKPs) effectively protect user data in social media and IoT applications without significantly degrading performance.

Blockchain offers privacy-enhancing features, such as encryption and Zero-Knowledge Proofs (ZKPs), which are hypothesized to protect sensitive user and IoT data without severely affecting the system's overall performance. By using cryptographic methods, user data can be secured from unauthorized access, and only authorized parties will have the keys to decrypt and view the data. Despite the potential computational overhead of these techniques, it is expected that they will offer robust privacy protection with minimal performance degradation.

H_5: Blockchain's decentralized architecture will reduce the risk of large-scale data breaches in social media platforms and IoT ecosystems.

In centralized systems, a single point of failure can lead to large-scale data breaches, particularly in platforms where user data is collected and stored in massive volumes, such as social media and IoT networks. Blockchain's decentralized architecture, which distributes data across multiple nodes, is hypothesized to mitigate this risk. Since data is not stored in a single repository, potential attackers will find it much more difficult to compromise the entire system, thus reducing the likelihood of widespread breaches.

8. Conclusion

The convergence of blockchain technology and big data offers a promising avenue for tackling the pressing challenges of data security, privacy, and integrity in diverse sectors such as social media, IoT, and industrial systems. This review has demonstrated that blockchain's decentralized, tamper-resistant architecture enhances the security of big data environments by ensuring transparent and secure data storage, enabling fine-grained access control, and safeguarding data integrity. The use of cryptographic techniques and smart contracts further strengthens privacy protection and ensures trust in data-sharing processes. However, the full potential of blockchain in big data applications is hindered by challenges such as scalability, high computational costs, and integration complexities. Addressing these limitations will require advancements in consensus mechanisms, improved scalability solutions, and better compatibility with existing big data infrastructures. As research and development progress, blockchain-enabled big data frameworks are poised to become a cornerstone in the future of secure, privacy-preserving, and efficient digital systems.

Reference

1) Ali, S., Fadlullah, Z. M., & Kato, N. (2020). Integrating Blockchain and Big Data for Social Media Security. IEEE Access, 8, 105674-105688.
2) Rahman, M.S., Chamikara, M.A.P., Khalil, I., & Bouras, A. (2022). Blockchain-of-Blockchains: An Interoperable Blockchain Platform for Ensuring IoT Data Integrity in Smart Cities. Journal of Industrial Information Integration, 30, 100408.
3) Juma, M., Alattar, F., & Touqan, B. (2023). Securing Big Data Integrity for Industrial IoT in Smart Manufacturing Based on Trusted Consortium Blockchain. IoT Journal, 4(1), 27-55.
4) Makhdoom, I., Zhou, I., Abolhasan, M., Lipman, J., & Ni, W. (2020). PrivySharing: A Blockchain-Based Framework for Privacy-Preserving and Secure Data Sharing in Smart Cities. Computers & Security, 88, 101653.
5) Liu, L., Li, J., Lv, J., Wang, J., Zhao, S., & Lu, Q. (2024). Privacy-Preserving and Secure Industrial Big Data Analytics: A Survey and the Research Framework. IEEE Internet of Things Journal.
6) Mitra, A., Bera, B., Das, A.K., Jamal, S.S., & You, I. (2023). Impact on Blockchain-based AI/ML-Enabled Big Data Analytics for Cognitive Internet of Things Environment. Computers & Communications, 197, 173–185.
7) Ali, S., Fadlullah, Z. M., & Kato, N. (2020). Integrating Blockchain and Big Data for Social Media Security. IEEE Access, 8, 105674-105688.
8) Roy, S., Salah, K., & Baggili, I. (2019). Blockchain for IoT Security and Privacy: The Case Study of Social Media. IEEE Internet of Things Journal, 6(3), 5340-5348.
9) Zhang, C., Wang, W., Li, P., & Fang, Y. (2021). Smart Social Media: A Blockchain-Based Decentralized Approach for Data Security and Privacy. Journal of Information Security and Applications, 60, 102850.
10) Al-Muhtadi, J., Alotaibi, A., & Siddiqui, S.T. (2021). Blockchain Technology for Privacy-Preserving Social Media. Journal of Network and Computer Applications, 177, 102899.