DATA TRUST FRAMEWORK USING BLOCKCHAIN
TECHNOLOGY ADAPTIVE TRANSACTION VALIDATION

1 MEDASANI GOPI
[2] Mr. Naga Srinivasa Rao

[1]PG STUDENT ,DEPT OF MCA

[2] **Asst. Prof**, Dept of MCA

SREE KONASEEMA BHANOJI RAMARS P.G. COLLEGE (AMALAPURAM)

ABSTRACT

Trust is the main barrier preventing widespread data sharing. The lack of transparent infrastructures for implementing data trust prevents many data owners from sharing their data and concerns data users regarding the quality of the shared data. Data trust is a paradigm that facilitates data sharing by forcing data users to be transparent about the process of sharing and reusing data. Blockchain technology proposes a distributed and transparent administration by employing multiple parties to maintain consensus on an immutable ledger. This paper presents an end-to-end framework for data trust to enhance trustworthy data sharing utilizing blockchain technology. The framework promotes data quality by assessing input data sets, effectively manages access control, and presents data provenance and activity monitoring. We introduce an assessment model that includes reputation, endorsement, and confidence factors to evaluate data quality. We also suggest an adaptive solution to determine the number of transaction validators based on the computed trust value. The proposed data trust framework addresses both data owners' and data users' concerns by ensuring the trustworthiness and quality of the data at origin and ethical and secure usage of the data at the end. A comprehensive experimental study indicates the presented system effectively handles a large number of transactions with low latency.

Keywords: Blockchain, Data Trust, Data Sharing, Distributed, Access Control.

I. INTRODUCTION

Data sharing has emerged as a significant concern when it comes to issues of privacy, confidentiality, data abuse, and legal and ethical violations. The absence of a transparent and reliable framework for data trust poses obstacles for many data owners who wish to share their data, which could be crucial for various research purposes. Data sharing is not only a major concern for data owners, but also for data users who want to ensure the trustworthiness and reliability of the data they receive. Data trust is a relatively new concept that aims to facilitate data sharing by requiring data users to be transparent about the process of sharing and reusing data. It encompasses legal, ethical, governance, organizational, and technical requirements to enable effective data sharing. Previous studies have proposed the use of web observatories and institutional repositories as potential means of implementing data trust.

Blockchain technology holds significant promise in establishing a practical data trust framework by revolutionizing current auditing practices and enabling the automatic enforcement of smart contracts. Blockchain's inherent properties make it well-suited for ensuring data integrity, transparency, and immutability. By leveraging blockchain technology, a robust and trustworthy environment for data sharing can be created.

Implementing a data trust framework would benefit both data owners and data users. For data owners, it would provide assurance that their data will be handled responsibly and used for its intended purpose. Data users, on the other hand, would have confidence in the quality and authenticity of the data they receive. Ultimately, establishing trust in the data sharing process is essential for promoting collaboration, innovation, and advancements in various fields of research.

In conclusion, addressing the concerns surrounding data sharing requires the development of a transparent and trustworthy framework for data trust. By incorporating elements such as legal compliance, ethical considerations, effective governance, organizational structure, and leveraging technologies like blockchain, we can create an environment that encourages responsible and secure data sharing while protecting privacy and confidentiality.

Numerous other research studies have explored the potential of blockchain in terms of sharing data, establishing trust, and controlling access. However, these studies are often fragmented and tend to concentrate on a specific step or aspect of data sharing or solely address the concerns of data owners. One possible application of blockchain is as a means of facilitating trust between data controllers and data users. The decentralized, secure, and dependable nature of blockchain technology can enhance the reliability and credibility of the data trust framework.

In our research, we aim to present a robust framework that utilizes blockchain technology to ensure data trust. This framework is designed to uphold the reliability and quality of data for both data users and owners, while also promoting ethical and secure data usage. To achieve this, we have devised a trust model that evaluates the trustworthiness of input data sets based on three key parameters: the endorsement and reputation of the data owner, the endorsement of the data asset, and the data owner's confidence level in the provided data set. These parameters are all stored on a ledger and are updated with each new transaction.

II. LITERATURE REVIEW

Shala et al. established a reward system to encourage IoT network peers with low trust scores to raise it. The motivational system makes use of control loops with a goal trust score. A bundle of incentives, such as discounts for other services, will be provided to service providers with low trust ratings to entice them to deliver a better service in return for the promised advantages. In, authors introduced an incentive-based strategy to motivate medical data owners to share their high-quality (actual and practical) data and receive income, as well as miners who profit by taking part and confirming transactions.

Wang et al. developed a system for an incentive that protects anonymity in order to generate high-quality crowdsensing contributions. Participants are encouraged by the trust mechanism to give their high-quality sensing data in exchange for Bitcoin or Monero. Data miners also make money by ensuring the accuracy of the data.

Zavolokina et al gave a financial incentive for joining the network and offers top-notch information for automobile dossiers. The system anticipates that by penalising bad behavior, mistakes would be reduced. For automatically calculating and implementing incentives, they use smart contracts. Blockchain technology and smart contracts were used by Shrestha and Vassileva to encourage data owners to contribute their research data without giving up ownership of it. In order to guarantee high-quality data exchange in the vehicular network, a subjective logic model has been employed to evaluate the reputation of nodes.

Dedeoglu et al. provided a trust model to evaluate the accuracy of data collected by IoT network sensor nodes. The credibility and repute of the data source, together with evidence from observations made by other neighboring sensor nodes, make up the model. Blockchain is also used to monitor the accuracy of shared data by looking for incorrect or suspect data that may have been gathered by IoT or mobile crowd sensing. Choudhury et al. maintained data privacy while ensuring data quality. As network members, regulatory bodies evaluate the accuracy of the data. By establishing activity-specific private channels, data privacy is protected. Delegated proof of reputation (DPoR), a lightweight consensus technique, was introduced by An et al. to address the challenging computing issue relevant to crowd sensing nodes' data quality management. Through the use of smart contract verification procedures.

Huang et al. made sure that the data gathered from sensor nodes in the crowdsensing network was of high quality. To promote the sharing of high-quality data, Su et al. created a two-tier incentive scheme based on reinforcement learning (RL). In the edge computing layer, Casado- Vara et al. also introduced a cooperative approach based on game theory to support data quality and false data detection

III. METHODOLOGY

In the proposed system, the system proposes an end-to-end framework for data trust based on blockchain, which ensures the trustworthiness and quality of the data at origin for data users and ethical and secure usage of data for data owners. First, we introduce a trust model to assess input data sets' trustworthiness using three parameters: data owner endorsement and reputation, data asset endorsement and data owner confidence level in the provided data set. All these parameters are recorded on the ledger, and they will be updated with every new transaction.

The methodology for developing the Data Trust Framework Using Blockchain involved a multi-faceted approach that integrated blockchain technology with data trust principles. The methodology presented in the base paper provides a comprehensive approach to developing a data trust framework using blockchain technology. By incorporating trust models, access control management, data asset validation, and adaptive transaction validation, the framework ensures the trustworthiness, security, and quality of shared data. The methodology presented in the base paper aims to address the concerns of both data owners and data users by ensuring the trustworthiness and quality of data at its origin and promoting ethical and secure usage of the data. The effectiveness of the proposed system is evaluated through a comprehensive experimental study, demonstrating its ability to handle a large number of transactions with low latency

The methodology encompasses various stages, including trust model development, access control management, and data asset validation.

Trust Model Development: The first step in the methodology is the development of a trust model to assess the quality and trustworthiness of input data sets. The trust model incorporates parameters such as data owner reputation, data asset endorsement, and data owner confidence level in the provided data set. These parameters are recorded on the blockchain ledger and updated with each new transaction. The trust model utilizes blockchain technology to calculate the trust value of data sets, ensuring that only trusted data sets are confirmed and recorded on the ledger.

Access Control Management: The next stage of the methodology focuses on implementing a secure and trustable access control management system using distributed ledger technology. Blockchain's features, including transparency, auditability, and trust distribution, are leveraged to achieve secure and fine-grained access control. Smart contracts are designed to handle access requests, consent management, and access permissions. Data asset owners have full control over their data assets and can regulate access permissions without relying on third parties. Access permissions can be granted to specific users or sub-groups of users belonging to one or multiple organization.

Data Asset Validation: To ensure the accuracy and quality of data assets, a validation process is implemented. Data sets with lower trust values are considered suspicious and require validation by multiple verifiers. This adaptive selection of verifiers strikes a balance between data asset quality and system performance. Data investigators are granted access to a small chunk of data to examine its accuracy and quality. Once the data sets are recorded as data assets on the ledger, data users interested in accessing a data set can prepare a request, which is directly received by the data owner. Using blockchain and smart contracts, all transactions are automatically enforced, eliminating the need for third-party involvement.

Data Trust Portal (DTP) Architecture: The Data Trust Portal (DTP) architecture is a crucial component of the data trust framework, providing a platform for secure discovery and sharing of data sets. The DTP architecture is inspired by web observatory infrastructure and is designed to facilitate the implementation of a secure discovery and sharing protocol using metadata about data properties and its provenance.

The DTP does not store the data itself; instead, it acts as a centralized platform where data owners can register their data sets and define access control policies. Data owners are responsible for the protection of their data and implement interface methods to provide access to authorized users. The DTP ensures that data sets are discoverable and accessible to eligible parties while maintaining data privacy and security.

The architecture of the DTP consists of three core layers:

Data Layer: This layer represents the actual data sets owned by data owners. Data owners hold and manage their data, ensuring its protection and privacy. The data layer includes mechanisms for data de-identification, anonymization, and other techniques to safeguard sensitive information.

Access Layer: The access layer provides the mechanism for authorized users to discover and access data sets. It includes standardized access protocols and interfaces, both centralized and peer-to-peer, to facilitate data sharing. Access control mechanisms are implemented to ensure that only authorized users can access the data sets based on the defined policies.

Process Layer: The process layer governs the overall workflow of data protection and usage. It includes well-defined data governance roles and processes to ensure effective data usage, sharing, and reusing. This layer controls data access through standardized risk assessments and tailors data to specific queries. It also ensures accountability and auditing of data usage, providing a transparent history of data access requests and usage.

The DTP architecture serves as a foundation for the data trust framework, enabling secure and controlled data sharing while maintaining data privacy and trust. It provides a structured and standardized approach to data stewardship, promoting responsible data usage and fostering trust among data owners and users.

By incorporating the Data Trust Portal (DTP) architecture into the methodology, the data trust framework gains a robust and scalable platform for secure data discovery, sharing, and access control. The DTP architecture enhances the overall functionality and effectiveness of the framework, ensuring the trustworthiness and integrity of shared data sets.

Experimental Study and Performance Evaluation: To assess the effectiveness and performance of the data trust framework, an experimental study is conducted. The study involves simulating various scenarios and evaluating the framework's response in terms of transaction processing time, resource utilization, and scalability. Performance metrics such as throughput, latency, and system response time are measured and analyzed. The experimental study provides valuable insights into the framework's capabilities, limitations, and areas for optimization.

Data Provenance and Activity Monitoring: To enhance transparency and accountability, the data trust framework incorporates data provenance and activity monitoring. Data provenance tracks the origin and history of data sets, providing a clear audit trail of data sources, transformations, and modifications. This information is recorded on the blockchain ledger, ensuring the integrity and traceability of data assets. Activity monitoring captures and logs all user interactions and system activities, allowing for real-time monitoring and detection of any suspicious or unauthorized actions.

Security and Privacy Measures: The data trust framework incorporates robust security and privacy measures to protect sensitive data and ensure compliance with data protection regulations. Encryption techniques are employed to secure data transmission and storage, preventing unauthorized access or tampering. Access control mechanisms, including authentication and authorization, are implemented to restrict data access to authorized users only. Privacy-enhancing technologies, such as differential privacy or anonymization techniques, may also be employed to safeguard individual data privacy.

Continuous Improvement and Maintenance: The data trust framework is not a one-time implementation but requires continuous improvement and maintenance. Feedback from users, stakeholders, and system monitoring is collected to identify areas for enhancement and optimization. Regular updates, bug fixes, and security patches are released to ensure the framework remains robust, secure, and up-to-date with evolving technologies and data governance practices.

IV. SYSTEM ARCHITECTURE

As we previously mentioned, the goal of our proposed strategy is to develop a framework for data trust that benefits data owners as well as users. In order to achieve this aim, our system architecture consists of two primary elements. 1) A traceable and secure access control system; 2) A trust model to assess the quality of incoming data sets. Our data trust framework design is shown in Figure. In terms of the input data sets, we model trust. Using an application based on blockchain technology, our system determines the trust value of any given initial data set.

The system only logs trusted data assets on the ledger and uses this value to guarantee that only trusted data sets are validated. It is necessary to have additional verifiers validate data sets that have lower trust values because they are deemed suspicious. The number of verifiers is adaptively chosen, offering a reasonable compromise between the system's performance and resource consumption and the quality of its data assets. Data investigators would have limited access to data in order to prevent a data breach. That little section is used to assess the quality and correctness of the data.

The data users who are interested in gaining access to a data set can prepare a request to access the data set once the information in the data set is documented as data assets on the ledger. Requests for access to their data sets will be sent directly to the data owner, who will then determine the terms and conditions for such access. All transactions are automatically enforced and involve no third parties when using blockchain and smart contracts. The resource owner might easily develop and store the access control policies on the blockchain through a smart contract.

The 0. illustrates the high-level process flow of this system. It emphasizes the secure storage of data in a blockchain, the generation of metadata, data integrity verification, and access control mechanisms to protect and control access to the stored data.

The flow chart and various entities in it are explained following to the figure

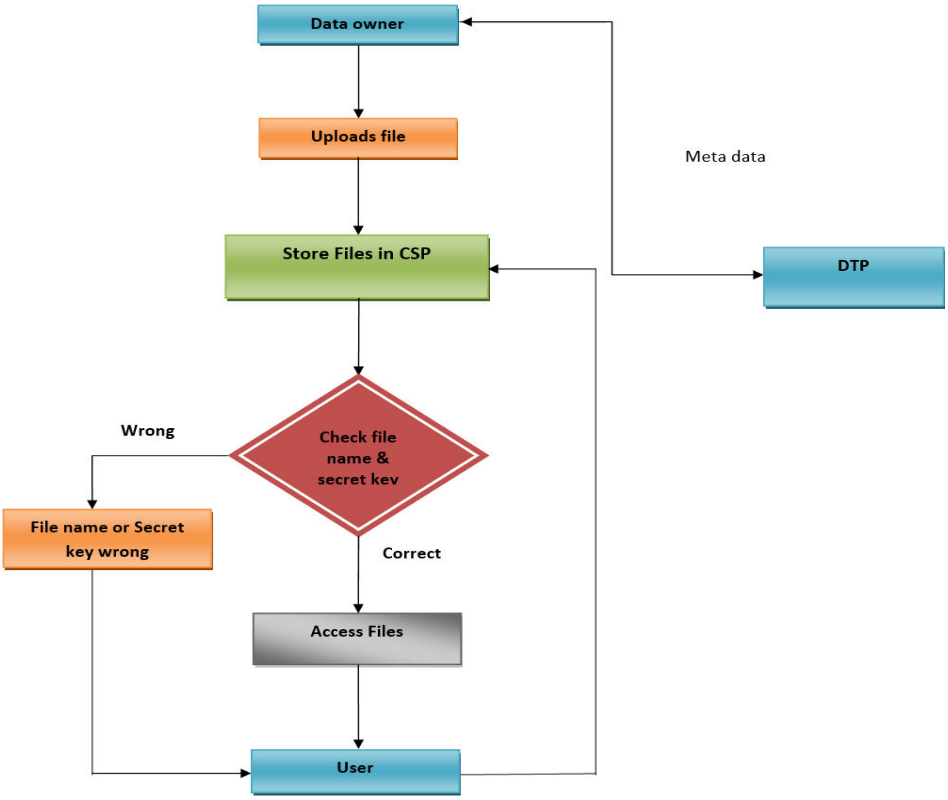


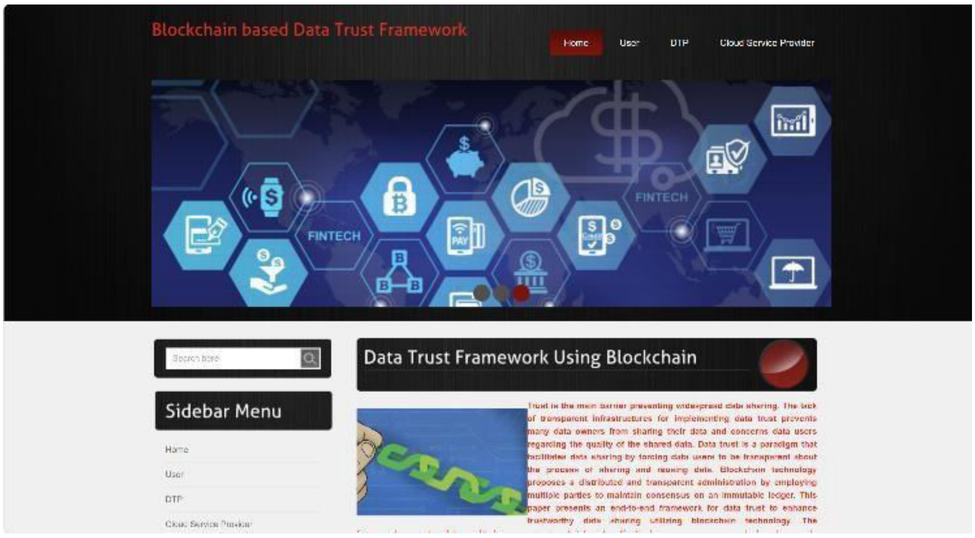
Fig 1: Flow chart

Fig 1. This is a figure caption. It appears directly underneath the figure.

1. **Data Owner:** The process begins with the data owner, who possesses the data that needs to be stored and secured using the blockchain-based data trust framework.
2. **Uploads File:** The data owner uploads the file or data to the system. This can be done through a user interface or an application provided by the framework.
3. **Store Files in Blockchain:** The uploaded file is stored in the blockchain. Unlike traditional storage systems, the blockchain provides a decentralized and immutable ledger where data can be securely stored.
4. **Generate Metadata:** Alongside the file storage, metadata is generated. Metadata includes information about the file, such as its name, size, type, timestamp, and any other relevant details that describe the data.

5. **Verify Data Integrity:** The framework performs a verification process to ensure the integrity of the data stored in the blockchain. This verification can involve cryptographic techniques like hashing or digital signatures to detect any tampering or unauthorized modifications.
6. **Access Control:** When a user requests access to the stored data, the framework initiates an access control mechanism. This mechanism verifies the user's identity and permissions to determine if they are authorized to access the data.
7. **Check User Credentials:** The user's credentials, such as username and password or cryptographic keys, are checked against the stored access control information. This step ensures that only authorized users can access the data.
8. **Correct/Wrong:** Based on the user's credentials, the system determines whether the access request is correct or wrong. If the credentials match the stored information, the access request is considered correct.
9. **Grant Access:** If the access request is correct, the framework grants the user access to the requested data. The user can then retrieve or perform operations on the data stored in the blockchain.
10. **Access Denied:** If the access request is wrong, indicating incorrect or insufficient credentials, the framework denies access to the data. This prevents unauthorized users from accessing sensitive information.
11. **User:** The user represents any entity or individual seeking access to the data stored in the blockchain. The framework ensures that only authorized users can access and interact with the data.

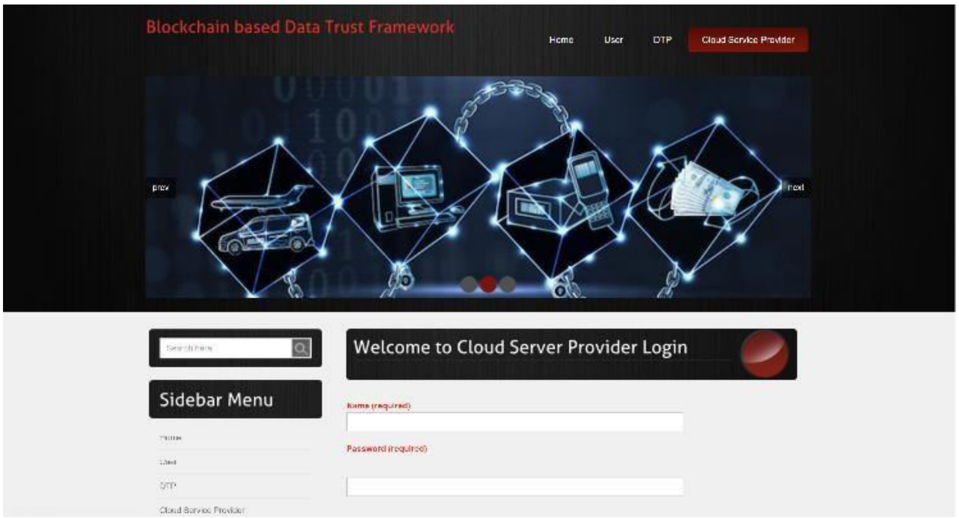
V. RESULT AND ANALYSIS



Data Trust Framework Using Blockchain

Trust is the main barrier preventing widespread data sharing. The lack of transparent infrastructures for implementing data trust prevents many data owners from sharing their data and concerns data users regarding the quality of the shared data. Data trust is a paradigm that facilitates data sharing by forcing data users to be transparent about the process of sharing and reusing data. Blockchain technology proposes a distributed and transparent administration by employing multiple parties to maintain consensus on an immutable ledger. This paper presents an end-to-end framework for data trust to enhance trustworthy data sharing utilizing blockchain technology. The

Fig 2: Home



Welcome to Cloud Server Provider Login

Name (required)

Password (required)

Fig 3: Cloud Server Provider login

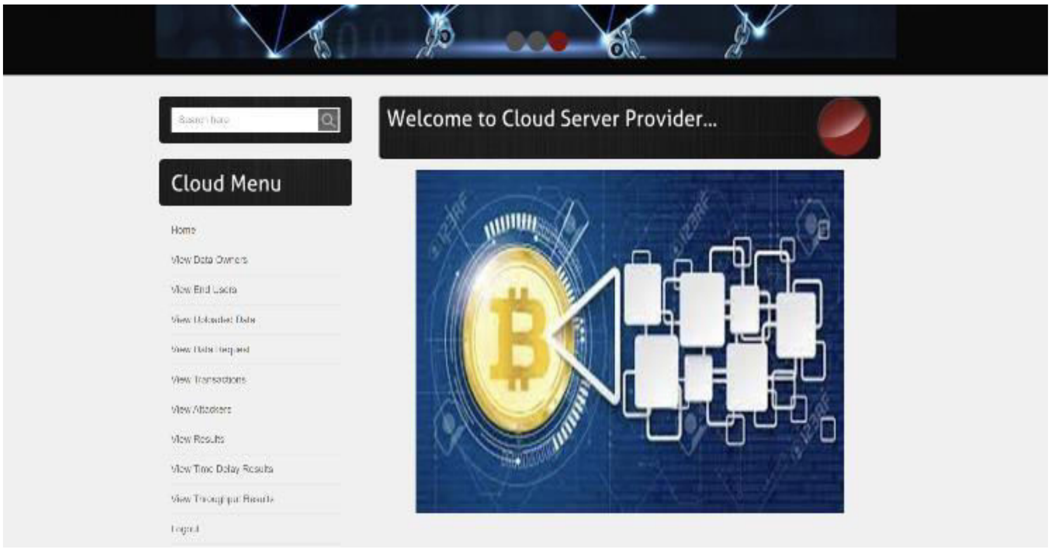


Fig 4: Welcome to CSP

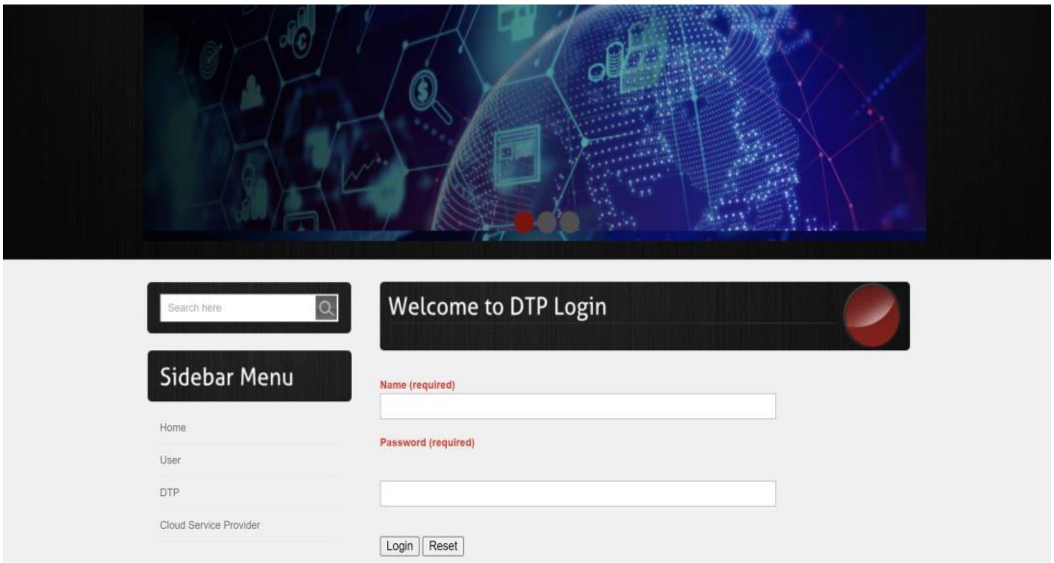


Fig 5: DTP login

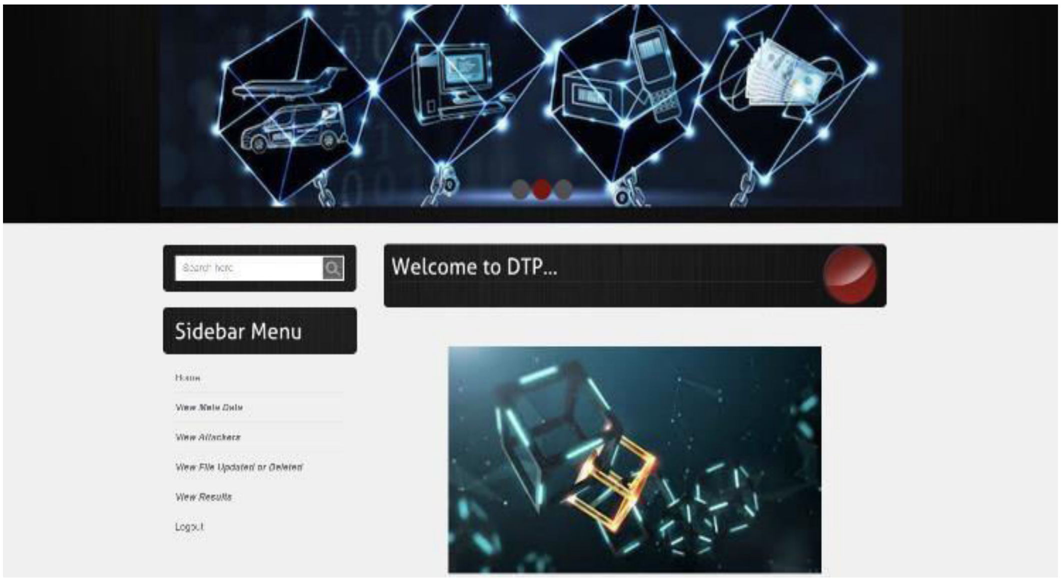


Fig 6: DTP

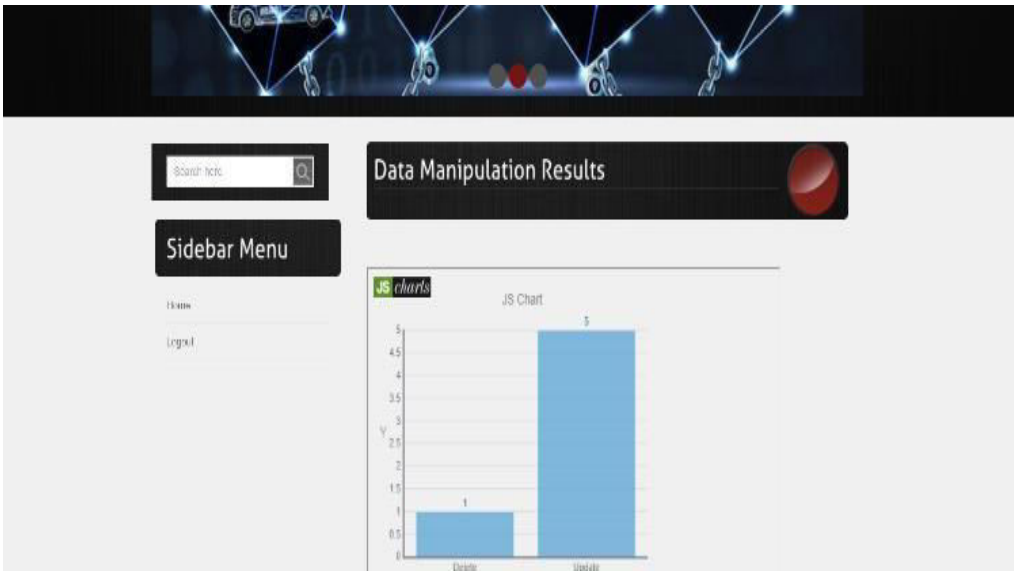
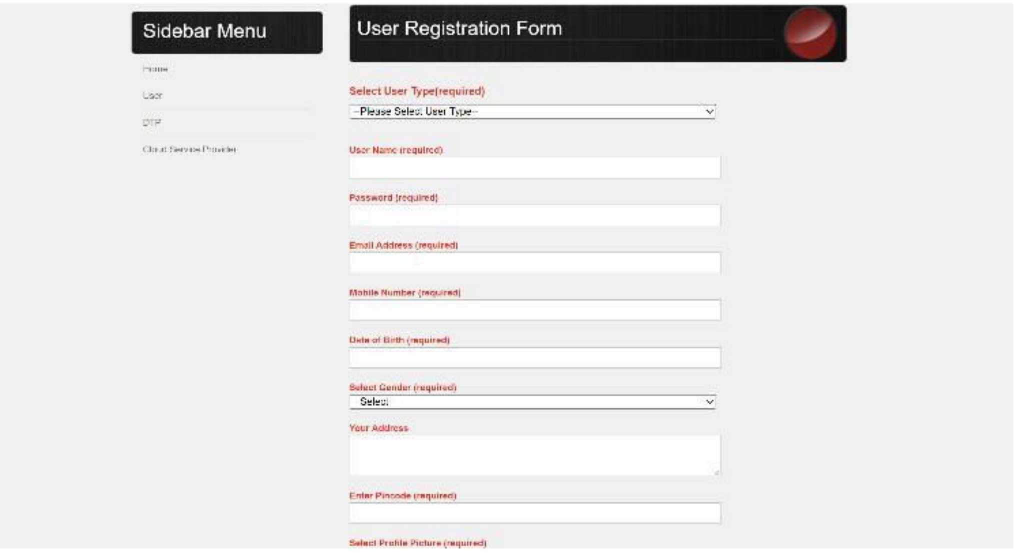


Fig 7: Manipulation Results



Select User Type(required)

--Please Select User Type--

User Name (required)

Password (required)

Email Address (required)

Mobile Number (required)

Date of Birth (required)

Select Gender (required)

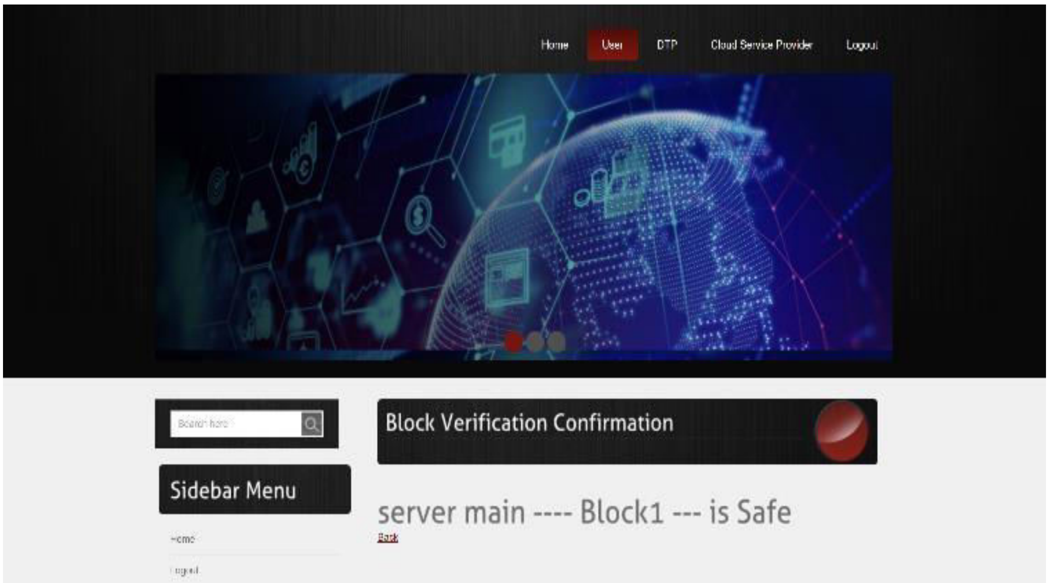
Select

Your Address

Enter Pincode (required)

Select Profile Picture (required)

Fig 8: Registration Form



server main ---- Block1 --- is Safe

Fig 9: Data Auditing

VI. CONCLUSION

Current systems are limited in providing a practical and transparent approach for data sharing due to the lack of trust in both parties. In this paper, we introduced an end-to-end data trust framework using permissioned blockchain. Our designed framework assesses the quality of input data using a novel trust model, including the data owner's reputation, endorsements and confidence in provided data. Therefore, the data users ensure that the available data set's quality has been adaptively examined and updated.

Data owners also can benefit from secure, transparent, and automatic access management using smart contracts. They have full control over their data assets, and they are the only actors in the system who can regulate access permissions without relying on third parties. By exploiting blockchain's provenance and audibility, data owners can monitor the trace of access regulations and modifications on their data assets.

Moreover, valuable logs can be extracted from the ledger to present a transparent view of the system, identify suspicious requests, and detect protocol breaches leading to discovering possible threats. Evaluation results indicate the system's effectiveness in handling a large number of transactions for writing, updating, and querying trust parameters value. As a future direction, we are looking toward improving the credibility of our framework by adding incentives to encourage honest participation of the users by adding endorsements and ratings. Moreover, identifying invalid assessments because of inputs from disruptive users is another important step to enhance the solution.

VII. REFERENCES

[1] K. O'hara, ``Data trusts: Ethics, architecture and governance for trustworthy data stewardship,'' Univ. Southampton, Southampton, U.K., Tech. Rep., 2019.

[2] A. Alsaad, K. O'Hara, and L. Carr, ``Institutional repositories as a data trust infrastructure,'' in Proc. Companion Publication 10th ACMConf.Web Sci., Jun. 2019, pp. 1_4.

[3] S. Rouhani and R. Deters, ``Security, performance, and applications of smart contracts: A systematic survey,'' IEEE Access, vol. 7, pp. 50759_50779, 2019.

[4] J.-H. Cho, K. Chan, and S. Adali, ``A survey on trust modeling,'' ACM Comput. Surv., vol. 48, no. 2, pp. 1_40, Nov. 2015. SPECIALUSIS UGDYMAS / SPECIAL EDUCATION 2022 2 (43) 1061.

[5] Z. Yan and S. Holtmanns, ``Trust modeling and management: From social trust to digital trust,'' in Computer Security, Privacy, and Politics: Current Issues, Challenges, and Solutions. Hershey, PA, USA: IGI Global, 2008, pp. 290_323.

[6] S. Stalla-Bourdillon, G. Thuermer, J. Walker, L. Carmichael, and E. Simperl, ``Data protection by design: Building the foundations of trustworthy data sharing,'' Data Policy, vol. 2, pp. 1_10, Jan. 2020