Genomics on the Chain: Exploring the Benefits of Integrating Genetic Data with Blockchain Technology

Amrutha.B K, Dr.B.Gomathy
Research Scholar, Department of ICT, Anna University, Chennai
Professor, Department of Computer science, PSG itech, Neelambur, Coimbatore

Abstract- **Genomics data is extremely important and sensitive. It is difficult to handle genomic data due the large quantity. People's genomic and health data must be combined and analysed in order to fully realise the potential of these technologies. Rapid generation of enormous quantities of genomic data is a consequence of the advancements in genome technology. Security measures that maintain confidentiality and restrict access are crucial since genetic data contains a variety of information that is sensitive to privacy. This may result in the creation of benefit sharing, data autonomous government, and better safety. Medical records are protected by a multitude of security measures. Nevertheless, there are also obstacles to the gathering and sharing of genetic data, including inadequate data quality, information islands, manipulation-induced distortions, insufficient documentation, leakage of personal data, and operations incorporating grey data. Moreover, problems with privacy arising from the sharing of genetic data have not yet been resolved. The blockchain platform offers novel opportunities for genetic data management and security. In this work, with blockchain technology is used to create a safe genetic data management system. and also analyses the example of from the perspectives related to information possession, data transferring, and data protection, LifeCODE.ai, a blockchain-based genetic big data platform, demonstrates how blockchain makes it feasible to store and manage genetic data. An estimated three billion base pairs make up the human genome. Even certain common diseases like diabetes, hypertension, and the like have a genetic sensitivity. certain genetic diseases are rare. As a result, processing datasets calls for specialised pipelines and equipment.**

Keywords: **Genetic, Blockchain, Genomic data**

I. INTRODUCTION

The rapid advancement of genomic as well as genetic technology, such as genome editing and next-generation sequencing, has led to significantly more precise and efficient disease treatment. The full potential of these technologies requires the combination and investigation into individuals inherited and healthcare information. There are numerous barriers to the collecting and distribution of genetic data, such as poor accuracy of the data, resource islands, distortions caused by manipulation, missing records, confidentiality breaches, and transactions involving grey data.

Blockchain technology has been suggested by a number of researchers as a solution to these issues. Blockchain is one sort of record form that makes sure the data is accurate. Initially, information about transactions for cryptocurrencies (like Bitcoin and Ethereum) was managed using it. Furthermore, blockchain has been applied to the management of recorded data histories, including medical records, in addition to the banking industry.

The main goal is to show how emerging blockchain technology's decentralisation, traceability, encryption, and antitampering features can be used to manage and secure confidential personalised health information. Blockchain-based genomic data LifeCODE.ai is used to demonstrate that blockchain makes it easier to keep and handle genetic information from the standpoints of data owning, exchange of information, and overall security.

A. Introduction to Genomics

The structure, development, mapping, and alteration of genomes make genomics studies in biology unique. The term "genome" refers to an organism's entire genetic code, or DNA. Genomic sequencing and analysis also include the assembly and examination of complete genomes' structure and function using high throughput DNA sequencing and bioinformatics.

Each gene datum is much larger than a typical healthcare data set, it takes a long time to sign, decrypt, and encrypt the full gene data set. Moreover, over 99.9% of the gene information is shared by all humans, negating the necessity for encryption or sign

the complete genome. Alternatively, by just safeguarding the DNA owner's share, the confidentiality and integrity of the data can be ensured at a lower computational cost. The expected increase in demand for gene data sharing is a further concern about gene data. While key management and access control schemes are reliable means of securely exchanging data, they are not a complete solution to the sharing problem. A user can request access to shared gene data, but once the data is transmitted, the user can keep it forever. If the user intentionally or unintentionally divulges their genetic data, the owner's privacy will be permanently breached.

B How Blockchain in Genomics?

Blockchain technology has a lot of potential for developing new systems and revolutionising the medical field.

All network information is recorded on a decentralised digital public ledger called a blockchain. Blockchain is being used extensively across a number of industries, including finance, healthcare, and education.

- Compared to traditional systems, blockchain platforms enable faster and more efficient data transactions.
- The main advantage of the blockchain is its immutability as opposed to the data in the central system, which is manipulated. Because the blockchain is a transparent system, every data transaction made through it becomes increasingly traceable and transparent.
- Blockchain technology is better at reducing fraudulent and unauthorised activity than earlier systems.

C. Benefits of using Blockchain technology in Genomics?

- Genetic Data Security: Given the sensitivity and significance of genetic data, blockchain technology provides better data security and integrity. Security solutions like encryption can help avoid data breaches, but they cannot provide complete protection. Numerous highly protected systems at major organisations are breached by hackers. However, blockchain technology helps companies by providing better protection against data breaches. Blockchain stores data securely using hashing algorithms, which benefits the organisation by facilitating data sharing and data security. Blockchain is being used extensively across a number of industries, including finance, healthcare, and education.
- Genome-wide Databases Sharing: Anonymous genetic information can now be shared worldwide through the use of genomic data and the network known as blockchain. Blockchain's decentralised structure makes it simple and safe for businesses to share data with one another. Data can be securely saved in a blockchain database by being placed in a unique ledger.
- Immutability of genomic data: Blockchain helps businesses safeguard information by enabling organisations' genetic data to remain unchanged. Since genetic data cannot be changed due to the decentralised nature of blockchain technology, any modifications made would be reflected on all nodes, making it impossible for anybody to cheat. It is therefore safe to say that sharing genetic data is possible.
- Efficiency: The company employs blockchain technology to improve system speed and efficiency by removing any third-party involvement that could cause errors while sharing genomic data. Data sharing thus gets simpler, more seamless, and quicker.
- Cost Reduction: Blockchain saves businesses money and fosters confidence among partners because it doesn't need a third party. Before blockchain technology, companies had to pay a third party a lot of money to maintain all the features that blockchain technology offered.

Due to the privacy-sensitive gene owner identification information contained in gene data, gene data privacy has drawn a lot of study interest. Numerous investigations have prevented the violation of gene owners' privacy through the use of membership inference attacks or attribute inference attacks. Numerous cryptography-based privacy-preserving solutions have been proposed to thwart these attacks. Nevertheless, there can be additional privacy risks if you solely use cryptography techniques. As previously mentioned, the scored gene data may be compromised for a number of reasons, including recipient negligence, if the recipient keeps the encrypted gene data forever, even in cases where the external recipient is trustworthy. We used the DP, which is used in many gene data management systems, to overcome this constraint.

Owing to DP's short history, not much research nor development has been done on gene data management systems based on DP. LDP has been used to safeguard data during storage, and previous

studies have shown statistical findings relevant to DP for gene data. To the finest of what we know, there is only one LDP-based method for gene data, and using the LDP-based approach causes additional challenges with efficiency.

The decentralised ledger known as "blockchain" (BC) records each and every transaction that takes place inside a network. Every transaction in the network's history is represented by an encrypted data unit called a block. Implementing decentralised technology will accelerate the shift from the existing hierarchical organisation to a decentralised, cooperative chain of command, enforce accountability, and call for increased security. A blockchain is utilised to preserve the data of a unique species of endangered animals by creating new offspring of those creatures that can be recognised by their identifying tags.

D. Application of Blockchain

Considering blockchain technology is still in its infancy, interest in it among researchers and practitioners is growing. In a nutshell, A blockchain is an openly viewable ledger that uses encrypted hash connections and timestamps to securely and permanently seal blocks. Its reliability, smart contract functionality, decentralisation, ability to resist manipulation, and encryption algorithm enable transaction activities without the requirement for a reliable outsider. Tracking asset ownership before, during, and after each transaction is safe and transparent thanks to a blockchain's usage of timestamp technology, which provides data traceability and verifiability. Second, the blockchain's free and open-source sharing protocol enables everyone to log and save data at the same time, ensuring that the transaction details are recorded and cannot be changed in the past without the network's approval. Lastly, a blockchain is very resilient and impervious to attacks and collusion because of its decentralised architecture and governance. Additionally, a blockchain can protect users' privacy by utilising cryptographic hash algorithms and asymmetric encryption, which enables users to encrypt data with their own private key. Recorded transactions cannot be changed in the past without network approval.

Last but not least, a blockchain's decentralised Structural and Administration make it very fault-tolerant and impervious to conspiratorial attacks. Additionally, a blockchain can safeguard users' privacy by employing asymmetric encryption, which lets individuals encrypt data with a private key, and hash functions that use cryptography.

II. BLOCKCHAIN TECHNOLOGY

APPLICATIONS IN BIOINFORMATICS AND HEALTHCARE

- Blockchain is a useful tool in bioinformatics and related sectors since it reduces the cost of analysis for genomics and health applications. Compared to traditional processes, blockchain technologies provide faster and more efficient data transactions.
- The primary benefit of the blockchain over the central system is its ability to prevent data manipulation. The blockchain's distributed architecture increases the visibility and traceability of every data exchange.
- Blockchain outperforms conventional systems in preventing fraud and unauthorised activities.

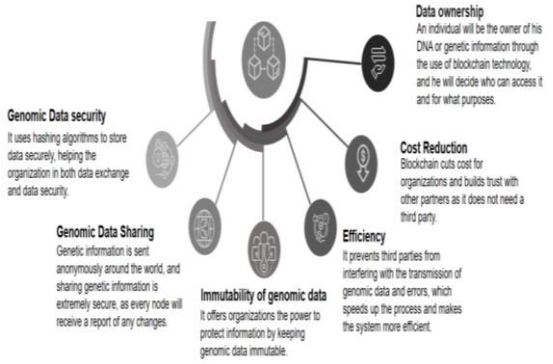

Fig 1: Advantages of Blockchain in healthcare

A.Data ownership is empowered by Traceability.

Traceability is the ability to use official identification records to verify the history, current location, or intended purpose of an item. Blockchain uses timestamp technology to fully record every step of the data creation and storage process, making the data traceable and verifiable. The record can be used to independently verify who owns the data, and as a result, each bit of data on the blockchain may be uniquely identified as coming from the specific data developer at the time of origination.

B Smart Contracts and Antitampering Strengthen Data Sharing

A collection of rules found in a smart contract facilitate automatic communication between the contract's parties. On the blockchain, no single node

may change any record. To ensure that smart contracts function, the antitampering functionality protects the data's originality and aids in implementing access. The negotiating and carrying out of an agreement or transaction are made easier, verified, and enforced by the smart contract code. By implementing these two blockchain characteristics, platform users will be encouraged to provide data for external benefits, improving the data sharing mechanism.

C Encryption algorithms and decentralisation enhance data security and privacy protection

Trust is built, security concerns are decreased, and data is managed with the use of decentralised storage. Secure electronic data transit includes the use of encryption algorithms. Data security and privacy are assured by the blockchain encryption technology, which also encrypts the data. In the case of online education, Sun et al. suggested using blockchain technology to safeguard data confidentiality and privacy as well as intellectual property.

D Blockchain Schemes for Genome Data

Large volumes of gene data are challenging to handle and store, and concerns remain over the security of this data. By 2025, a storage capacity of approximately 40 EB will be required since sequencing more genotypes would necessitate storage space, and a single human genome requires 100 GB of storage. And also, people's interest in their individual genetic information is growing as the number of genetic tests with direct-to-consumer findings rises. It is crucial to consider if these data are securely safeguarded both during use and storage.

Blockchain is predicted by experts to be essential for industries including data sharing, trust, and openness. Blockchain technology, which is based on a distributed system as opposed to earlier centralised systems, is extensively employed to handle security and privacy concerns. Blockchain technology guarantees integrity and allows for a decentralised ecosystem. Privacy concerns can be secured with the use of consensus techniques, hybrid signatures, and node authentication. It is suggested using blockchain technology to maintain patient data securely and analyse large amounts of medical data.

A few researchers have investigated the potential of blockchain technology in this field and have put forth a blockchain-based system that offers phenotypical expression mistakes reversal, genomic privacy, security, and anonymous data analysis. Moreover, companies like CrypDist, Nebula Genomics, and Gene-Chain are utilising blockchain technology to enhance the exchange of genetic data. making corrections to phenotypical expression errors.

III. PROPOSED SYSTEM MODEL

The proposed method, seen in Figure 1, combines two different types of storage, private and semi-private, with a blockchain system featuring a decentralised application (DApp). Both storage solutions retain the gene data, but to protect security and privacy, portions holding the owner's information are obscured by encrypting the gene data or by adding noise. Gene data that contain noise are kept in semi-private storage, while encrypted gene data are kept in private storage. Furthermore, the gene data owner can utilise DApp to monitor the usage and transmission of their gene data. Encrypted gene data are stored in private storage, whereas noise-containing gene data are stored in semi-private storage. Similarly, the owner of the gene data may utilise DApp to keep an eye on how their gene data is being used and sent.

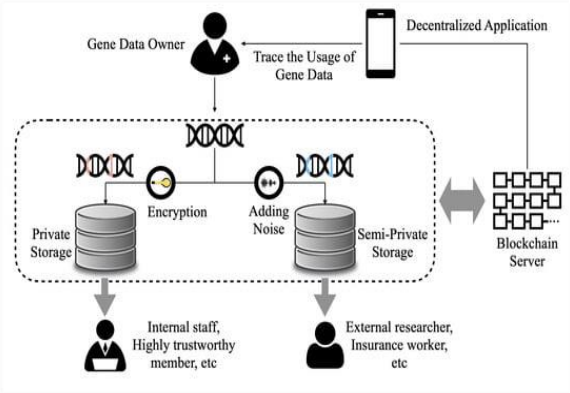

Fig 2: System Model

Private storage is only accessible to internal staff members and incredibly trustworthy members; only these people are able to unlock the encrypted sections and access and use the original genome data. It is presumed that these users don't give unreliable users the original data. Users would receive the actual gene data to examine.

Next, is regarding semi-private storage.

All genetic data, including noise, is stored in semi-private storage and is not recoverable from the original gene data by any other end users—apart from a very limited group of very untrustworthy individuals, such academics and insurance experts

employed by other companies. Data stored in semi-private storage is not available to the general public and can only be accessed by permitted individuals. Irreversible gene data are available, nonetheless, as these users might divulge stored gene data to other people or organisations, and after that, they generated noise using the LDP, adding the noise data to the gene data part that is private.

This two-storage-system approach lessens the chance that confidential material containing gene data may leak. Employees of external organisations or rather unreliable users are only allowed access to gene data with noise in the semi-private storage. Reputable internal researchers have access to encrypted gene data that is stored privately. Therefore, it is not feasible for authorised employees to depart from the recommended management method; they are the only ones who can access and retain the original gene data. The gene data that are not able to be converted back into the original form and contain noise are the data that can leave a system. Blockchain technologies not only provide encryption and noise production, but also ensure the anonymity of gene owners. The file size, hash value, creation/alteration time, and secured data sections are just a few of the several gene data-related details that are contained in each block. For access control purposes, lists of users with access to both encrypted and noise-encrypted gene data are maintained. Therefore, using smart contracts, the blockchain might be able to regulate access requests. Put another way, the system confirms that the blockchain system is accessible when it gets a request for gene data. If the blockchain smart contract responds with "accept," the proposed system forwards the required information.

Additionally, the DApp is linked to the blockchain to track the use and transfer of gene data. Furthermore, it suggests that the owner of the gene data may confirm who has access to it and how the user is using it on their mobile device. The suggested paradigm protects the privacy of the data owner by preventing outside parties from accessing the sensitive portions of the gene data. Additionally, only authorised individuals are permitted access to the storage via the blockchain's smart contract, and only trustworthy employees are able to retrieve original content from private storage.

Since the sensitive parts of the gene data won't be accessible to third parties, the proposed paradigm safeguards the privacy of the data owner. Additionally, only authorised individuals are permitted access to the storage via the blockchain's smart contract, and only trustworthy insiders are able to retrieve original content from private storage.

A. Blockchain-Powered Gene Data Administration

Figure 3 shows how the blockchain and the gene data storage interact with each other. In this part, the blockchain in relation to the mechanism for managing gene data that is being suggested. The two primary uses of blockchain are for access control and DNA data integrity verification. Here, we offer blockchain-based approaches to gene data management.

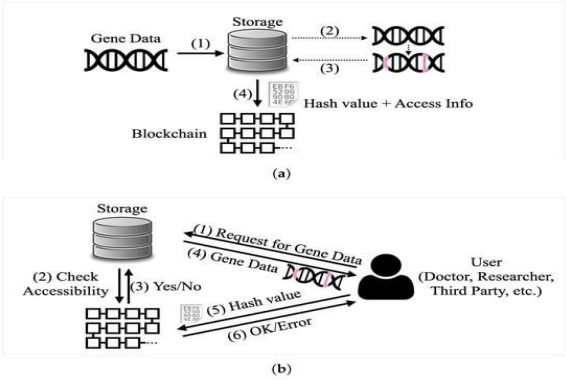

Figure3. Blockchain and storage interactions. Gene data can be (a) saved or(b) retrieved

Figure 3a: After the sensitive parts have been encrypted or have noise added to them. The gene data is sent encrypted to the data store. The hash function is then used to calculate the digest of the gene data. The purpose of this computation was to confirm data integrity. Access information, which identifies the member who can receive the material, is provided in addition to the hash value. After they are transmitted, these data are added to the ledger and a new block is made and added to the blockchain.

In Figure 3b :The data extraction procedure is explained. The store verifies that the gene data are accessible upon request from an internal or external user using a smart contract embedded in the blockchain. Data are only given to authorised users since the blockchain maintains access information about the data, and because of the blockchain's security, this information is never altered. The receiver runs the data's hash function and generates a message digest after data transfer. It then sends the hash value to the blockchain system to confirm the accuracy of the gene data.

B Current Developments in Genomics-Related Blockchain Systems

To better comprehend the trends of today, let's examine the period of blockchain-based technologies. Three periods of blockchain-based systems in genomics are distinguished by the NLM (National Library of Medicine) paper:

- The Proof of Concept Period (2016–2018):

Core development is needed for the suggested application because the blockchain concept is fresh in this era of technology. Developing a workable application to demonstrate the blockchain's suitability for health management systems is the primary objective of this period. Blockchain platforms in this period contain the following features:

Incorporate EHR and genetic data exchange, although research on EHR sharing was limited in previous eras. The specific design of the blockchain-based healthcare system was developed in this day and age, even though the recommended strategy steadily reduced the amount of blockchain and integrated other technologies into the system.

- Blockchain Development Era (2019–2020):

The majority of planned applications in this era are sophisticated cloud- and encryption-based ones. The main areas of research at this time were: Research focused on creating a blockchain-based health care framework that is comparable to the first era.

Research focusses on implementing blockchain-enabled modules, evaluating performance, and using various cryptographic techniques to increase system security rather than on disease prediction and patient monitoring methods.

- The Era of Blockchain as a Platform (From 2021 to the Present):

In this phase, blockchain is transformed into a platform that is home to extra AI-based algorithms. This time frame can be viewed as the beginning of creating a blockchain-powered data ecosystem. The main focus of this research was on creating a blockchain-based healthcare system that could monitor patients and anticipate diseases.

Artificial intelligence techniques were incorporated into the system to give these concerns more attention and assess the system's performance from this standpoint.

IV. FLOW DIAGRAM OF BLOCKCHAIN BASED GENOMIC DATA

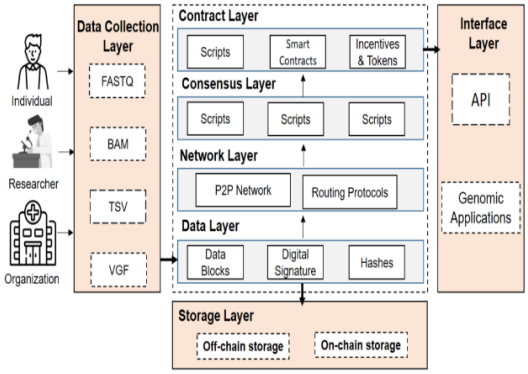

Fig 4: Flow diagram of blockchain based genomic data

Genomic data arrive in a variety of forms, such as BAM or FASTQ, and are sorted by each node in the first layer of architecture. These nodes in the system stand for a person, a team of researchers, or an institution that wishes to exchange genetic information. After collection, the data is transferred to the data layer, where blockchain features, like as cryptographic hashing and digital signatures, are used to protect the integrity and confidentiality of genetic data. Then, it is forwarded to the next level , for storage.

Storage layer is used to store data in a variety of ways based on the needs of either off-chain or on-chain storage. We can store data on-chain and essentially integrate it into the chain by appending the data (in binary format) to the transaction. Since the transaction will eventually be propagated to each node in the network, the data will finally become extremely accessible and immutable. It is essential to address the secrecy of the stored data because every node in the network has public access to on-chain data.

On-chain storage works best for Storage that is impervious to tampering is necessary for small data kinds. It also contains little genomic data and meta-data which is preserved on-chain are small data types like audit trails and observations of gene-drug interactions. Files like BAM or VCF include genomic data that is challenging to store on-chain.

Large data files or data requiring stringent access control are stored off-chain. In these situations, decentralised file systems like IPFS or cloud storage areutilised.

Hashing data yields a short string, which is used in off-chain storage methods. A smart contract or blockchain transaction can both effectively store this string. Next, a centralised or decentralised storage system is used to store the real data.

Next, the data is disseminated throughout the network via a designated network protocol, like P2P and Routing Protocol. At the consensus layer, the network's nodes determine the blockchain's current state by applying the consensus-building technique like Proof-of-Work.

Smart contracts are written at the contract layer. implemented in addition to support other application functionalities. Blockchain and smart contract interaction happens at the presentation layer.

V. CASE STUDY METHOD

The case study research approach can boost the evidence's efficacy and has a practical aspect . Ultimately, the case study approach is very descriptive, clearly analyses events, provides a strong feeling of reality, and is simple for readers to understand.

There were restrictions on how a case might be chosen in light of the research topic. Initially, the example organisation had to be a medical biomedicine medical firm. Secondly, in order to handle concerns with genetic data security and administration, the case organisation had to implement a blockchain platform. Finally, the organisation advocating for the case has to be open to receiving surveys on a regular basis.

The study's focus is LifeCODE.ai, a big-data platform enabled by blockchain for genomics. During the 2018 TechCrunch International Innovation Summit, LifeCODE.ai was initially made available. It builds a blockchain platform using a decentralised consensus technique to compile dispersed health data. LifeCODE.ai aims to enhance overall health outcomes through genomics research. It accomplishes this by creating a decentralised, transparent, and secure personal health data centre that makes it difficult to distinguish between health information owned by individuals, hospitals, doctors, R&D facilities for pharmaceutical corporations, and patients. Simultaneously, Laiyin Health, Individual users can now use a blockchain-based DApp developed by LifeCODE.ai that uses blockchain technology for data independence, token mechanism development, and the algorithm of encryption application. Unlike traditional apps, which have their backend code executed on centralised servers, DApps execute their backend code on a decentralised peer-to-peer network.

After that, on-site data were gathered through two interviews with the firm that created and oversees LifeCODE.ai, as well as interviews with departmental managers.

During these conversations, topics included the benefits of blockchain technology in 2018 and the functioning of LifeCODE.ai. Each interview, which lasted between sixty and ninety minutes, featured department managers and corporate officials from LifeCODE.ai. Every interview was captured on digital media and then transcribed. Simultaneously, we meticulously gathered secondary information about this platform from many publications, such as newspapers, periodicals, books, and the internet.

Blockchain as a service, which makes using it as simple as using the internet, is promoted by LifeCODE.ai. The following 3 product services are the major ones offered by LifeCODE.ai.

- Smart contracts: It disentangles blockchain-based, decentralised smart contracts.
- Bookkeeping function: Distributed bookkeeping is used In order to resolve the uncertainty. Without a central structure, everyone has access to the same book, guaranteeing a fair and open transaction process.
- Computational processing is responsible for resource management on platforms, data interpretation from gene sequencing, and data usability.

A Platform Architecture

By combining genetic and phenotypic health data with the best possible level of data privacy and quality, LifeCODE.ai builds a platform for data interchange, interoperability, and a plethora of services 1 for all participants. Figure 4 shows how the platform 1 layer and data layer are designed. The blockchain, infrastructure, business services, and interface layers are the four main levels that make up the application programming interface (API) and its associated customers. To increase data security, the data layer has trusted data storage and searchable encryption. The infrastructure layer comprises fundamental services among other things. Among other things, the security module has an access control network and basic security services.

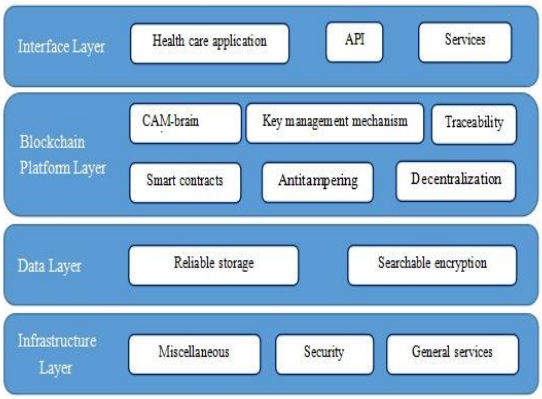

Fig 5: Architecture of LifeCODE.ai platform

The individual is the exclusive owner of the gene data in LifeCODE.ai. Personal genetic and phenotypic health data uploaded by users to the platform is often encrypted by default, preventing business participants from accessing it without the owner's consent. This method works similarly to how a bank would store and manage data on the platform.

B. Token mechanism in LifeCODE.ai

Data trading and sharing are made possible by LifeCODE.ai's token technology within a platform-based closed-loop ecosystem. The primary players in the LifeCODE ecosystem are individuals, medical centres, research institutes, insurance companies, healthcare providers, and pharmaceutical companies. In addition to effectively utilising the smart contract and antitampering features of the blockchain, LifeCODE.ai also leverages a token system to facilitate data sharing. Four stages are involved in the token mechanism.

- Using the genetic to import data owned by the data owner, the first step is to use ordered relational (GOR) architecture, which is specifically designed for genomic data storage, query, search, indexing, and many more analytics.
- Secondly, utilising the data are placed into a searchable encryption database after being either entirely or partially encrypted in accordance with the requirements using homomorphic encryption, tag-based fingerprint extraction, and asymmetric encryption techniques. In this instance, the data yield a label that matches and acts as an index for keyword searches.
- Third, LifeCODE.ai gathers, categorises, incorporates, and analyses the phenotypic data. LifeCODE.ai handles and integrates these data in a number of ways according to their unpredictable nature.
- The sharing mechanism's final technical support layer is the secure transfer of data during the data transaction process.

LifeCODE.ai exports health data that can only be searched for purposes of study and data analysis in an intermediate encrypted searchable data repository. Furthermore, only information that cannot be used against the data owner may be located using the search function, and the private section of the data is encrypted using the same method as the original data. Next, utilising zero knowledge proofs, the LifeCODE.ai platform cryptographically hides the genetic assets of a transaction. In this case, party A can demonstrate to party B that he is aware of a particular fact without really disclosing it, making party B the verifier and party A the prover.

VI. BLOCKCHAIN EMPOWERMENT PROCESS

The four stages of the blockchain activation process are as follows:

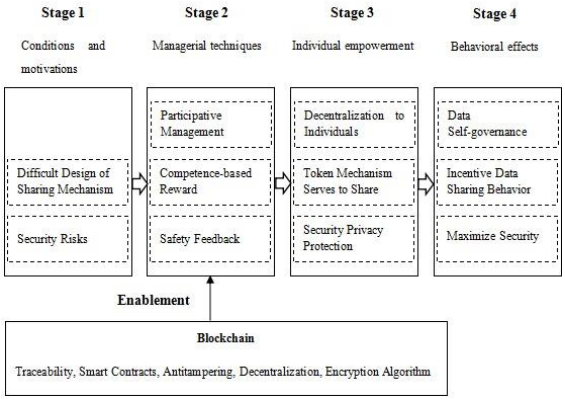

Fig 6: Blockchain empowerment Process

- Stage1:Condition and Motivation

Establishing genomic big data platforms is hampered by issues like uncertain data ownership, challenging sharing mechanism design, and security risk.

- Stage 2:Managerial Techniques

The unique characteristics of blockchain, such as its antitampering, decentralisation, traceability, smart contracts, encryption algorithm, and decentralisation, enable users of the blockchain platform to engage in data management. Additionally, they can receive token rewards through data transactions. One way to think of the token incentives is as a form of competency-based reward. Furthermore, a greater degree of data confidentiality and privacy can be guaranteed through the use of blockchain technology.

- Stage 3:Individual Empowerment

Blockchain technology offers fresh perspectives on how to build a big data platform for genomics. The use of blockchain technology gives people control over their data and gives them ownership of it. The data-sharing procedure can benefit from the token method, which also improves data security and privacy protection.

- Stage 4: Behavioural effects

In the end, the aforementioned elements will cause the users of the platform to behave differently. Data autonomy will be made possible by individual data ownership. More proactive sharing between individuals and businesses will be fostered by token systems. Furthermore, if security and privacy are consistently maintained, users will use this platform more frequently.

VII. CONCLUSION

Blockchain technology holds great promise for developing future systems that will reshape the medical industry. Blockchain technology in genomics is still in its infancy as applied by researchers and industry specialists. Many areas that might alter the current genetic ecology are still unexplored by them. Blockchain will also improve data sharing, process automation, and trust. Furthermore, blockchain provides incentives for design that will enable equitable sharing, storing, and processing of human genomic data. In brief, new avenues for genetic data management and security are made possible by the blockchain platform. Additionally, it can help people become more psychologically empowered, which will eventually result in the realisation of incentive-sharing, data self-governance, and security improvement. The most recent advancements in blockchain technology for healthcare were examined in this paper. Because sensitive data is handled and regulated, blockchain technology offers a decentralised network and is thought to hold great potential for use in the healthcare industry, especially for safeguarded DNA storage.

Determining the present state of blockchain research and its possible uses in genome sequencing and healthcare was the study's main goal. It's critical to keep in mind that different countries have distinct legal systems and that different cultures have varied views on who owns data. For this reason, the topic of data ownership requires more research. Transferring ownership of data to the people who produce it makes more sense than waiting for specific legal advice, especially when the debate in this paper is based on Chinese culture.

VIII. FUTURE SCOPE

Further investigation is necessary to explore the viewpoints of healthcare institutions and individual users about blockchain-powered data storage and management solutions. There are still certain concerns with blockchain technology, like capacity and regulatory constraints, that need to be resolved. Therefore, a comprehensive analysis including several scenarios and perspectives, together with extended observation, are imperative for the examination of blockchain in the context of gene data. Research and applications of blockchain technology are expanding in the medical field. According to current trends in blockchain research, the healthcare industry most frequently uses blockchain technology for data interchange, access control, and health records; supply chain management and prescription medication orders, on the other hand, are much less common. So, there is still a lot of untapped potential in blockchain. There is still room for research and the discovery of new applications because blockchain technology is still relatively new in the healthcare industry. In summary, blockchain technology should be used wherever it makes sense and is necessary. A comprehensive analysis from a range of perspectives and scenarios, along with continuous observation, are needed in order to examine blockchain in relation to gene data management and security assurance.

REFERENCES

[1] Sharp SA, Weedon MN, Hagopian WA, Oram RA. Clinical and research uses of genetic risk scores in type 1 diabetes. Curr Opin Genet Dev. 2018 Jun;50:96–102. doi: 10.1016/j.gde.2018.03.009. http://europepmc.org/abstract/MED/29702327. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[2] Ogishima S. [Human genome data and drug development] Gan To Kagaku Ryoho. 2018 Apr;45(4):597–600. [PubMed] [Google Scholar]

[3] Cai YD, Huang T. Accelerating precision medicine through genetic and genomic big data analysis. Biochim Biophys Acta Mol Basis

Dis. 2018 Jun;1864(6 Pt B):2215–7. doi: 10.1016/j.bbadis.2018.03.012. https://linkinghub.elsevier.com/retrieve/pii/S0925-4439(18)30092-9. [PubMed] [CrossRef] [Google Scholar]

[4] Low SK, Zembutsu H, Nakamura Y. Breast cancer: the translation of big genomic data to cancer precision medicine. Cancer Sci. 2018 Mar;109(3):497–506. doi: 10.1111/cas.13463. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[5] Interuniversity Institute of Bioinformatics in Brussels: IBsquare. 2016. [2019-01-30]. BRiDGEIris: BRussels Big Data Platform for Sharing and Discovery in Clinical GEnomics https://ibsquare.be/drupal/research/projects/bridgeiris-brussels-big-data-platform-sharing-and-discovery-clinical-genomics.

[6] Siwicki B. Healthcare IT News. 2018. [2019-01-30]. New Genomics Analytics Platform From Databricks Aims to Speed Discovery of New Treatments https://www.healthcareitnews.com/news/new-genomics-analytics-platform-databricks-aims-speed-discovery-new-treatments.

[7] Bhuvaneshwar K, Belouali A, Singh V, Johnson RM, Song L, Alaoui A, Harris MA, Clarke R, Weiner LM, Gusev Y, Madhavan S. G-DOC Plus - an integrative bioinformatics platform for precision medicine. BMC Bioinformatics. 2016 Apr 30;17(1):193. doi: 10.1186/s12859-016-1010-0. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1010-0. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[8] Li JQ. Open Medical Big Data and Open Consent and Their Impact on Privacy. Proceedings of the International Congress on Big Data; BigData Congress'17; June 25-30, 2017; Honolulu, HI, USA. 2017. [CrossRef] [Google Scholar]

[9] Pizzolante R, Castiglione A, Carpentreri B, de Santis A, Palmeiri F, Castiglione A. On the protection of consumer genomic data in the internet of living things. Comput Secur. 2018;74:384–400. doi: 10.1016/j.cose.2017.06.003. doi: 10.1016/j.cose.2017.06.003. [CrossRef] [CrossRef] [Google Scholar]

[10] Tao JL, Chen SY. Research on Personal Privacy Protection in Medical Big Data. Proceedings of the 2nd International Conference on Artificial Intelligence and Engineering Applications; AIEA'17; September 23-24, 2017; Guilin,China. 2017. pp. 1033–9. [CrossRef] [Google Scholar]

[11] Kulynych J, Greely HT. Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. J Law Biosci. 2017 Apr;4(1):94132.doi: 10.1093/jlb/lsw061. http://europepmc.org/abstract/MED/28852559. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[12] Balsari S, Fortenko A, Blaya JA, Gropper A, Jayaram M, Matthan R, Sahasranam R, Shankar M, Sarbadhikari SN, Bierer BE, Mandl KD, Mehendale S, Khanna T. Reimagining health data exchange: an application programming interface-enabled roadmap for India.

[13] JMedInternet Res. 2018Jul13;20(7):e10725 .doi: 10.2196/10725. https://www.jmir.org/2018/7/e10725/ [PMC free article] [PubMed] [CrossRef] [Google Scholar]G. Wood, "Ethereum: a secure decentralised generalised transaction ledger (2014)." 2017.

[14] M. Swan, Blockchain: Blueprint for a new economy. " O'Reilly Media, Inc.," 2015.

[15] M. Dhawan, "Analyzing safety of smart contracts," in Proceedings of the Conference: Network and Distributed System Security Symposium, San Diego, CA, USA, 2017, pp.16–17.

[16] N. Popper, "Understanding Ethereum, Bitcoin's virtual cousin," New York Times, vol. 1, p. 2017, 2017.

[17] N. Szabo, "Smart contracts http://www. fon. hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature," LOTwinterschool2006/szabo. best. vwh. net/smart.Contract. html Go to Ref. Artic., 1994.

[18] W. Zou et al., "Smart Contract Development: Challenges and Opportunities," IEEE Trans. Softw. Eng., vol. 47,no. 10, pp. 2084–2106, 2021, doi: 10.1109/TSE.2019.2942301.

[19] Q. Xu, Z. He, Z. Li, and M. Xiao, "Building an ethereum-based decentralized smart home system," in 2018IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), 2018, pp. 1004–1009.

[20] C. Dannen, Introducing Ethereum and solidity, vol. 1. Springer, 2017.

[21] P. Zhang, J. White, D. C. Schmidt, and G. Lenz, "Design of blockchain-based apps using familiar software patterns with a healthcare focus," in Proceedings of the 24th Conference on Pattern Languages of Programs,2017, pp. 1–14.

[22] "Remix - Ethereum IDE." https://remix.ethereum.org/ (accessed Dec. 19, 2022).R. Taş and Ö. Ö. Tanrıöver, "Building a decentralized application on the Ethereum blockchain," in 2019 3rdInternational Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2019, pp. 1–4.

[23] V. Buterin, "A next-generation smart contract and decentralized application platform," white Pap., vol. 3, no.37, pp. 1–2, 2014.

[24] K. Wu, Y. Ma, G. Huang, and X. Liu, "A first look at blockchain-based decentralized applications," Softw. -Pract. Exp., vol. 51, no. 10, pp. 2033–2050, 2021, doi: 10.1002/spe.2751.

[25] J. Angelis and E. R. Da Silva, "Blockchain adoption: A value driver perspective," Bus. Horiz., vol. 62, no. 3,pp. 307–314, 2019.

[26] R. Z. Farahani and M. Elahipanah, "A genetic algorithm to optimize the total cost and service level for just-in-time distribution in a supply chain," Int. J. Prod. Econ., vol. 111, no. 2, pp. 229–243, 2008.

[27] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," Multimed. Tools Appl., vol. 80, no. 5, pp. 8091–8126, 2021.

[28] "CryptoKitties | Collect and breed digital cats!" https://www.cryptokitties.co/ (accessed Dec. 19, 2022).

[29] F. Blum, B. Severin, M. Hettmer, P. Huckinghaus, and V. Gruhn, "Building Hybrid DApps using Blockchain Tactics -The Meta-Transaction Example," IEEE Int. Conf. Blockchain Cryptocurrency, ICBC 2020, 2020, doi:10.1109/ICBC48266.2020.9169423.

[30] A. Bogner, M. Chanson, and A. Meeuw, "A decentralised sharing app running a smart contract on the Ethereum blockchain," ACM Int. Conf. Proceeding Ser., vol. 07-09-Nove, pp. 177–178, 2016, doi:10.1145/2991561.2998465.Abdul-Sada et al., Al-Salam Journal for Engineering and Technology Vol. 2 No. 2 (2023) p. 37-47

[31] W. J. Buchanan, Blockchain and Crypto-currency. 2022. doi: 10.1201/9781003337751-11.

[32] T. Min and W. Cai, "Portrait of decentralized application users: an overview based on large-scale Ethereum data," CCF Trans. Pervasive Comput. Interact., vol. 4, no. 2, pp. 124–141, 2022, doi: 10.1007/s42486-022-00094-6.

[33] Z. Wang, H. Jin, W. Dai, K. K. R. Choo, and D. Zou, "Ethereum smart contract security research: survey and future research opportunities," Front. Comput. Sci., vol. 15, no. 2, 2021, doi: 10.1007/s11704-020-9284-9.