IRSTI 20.53.19 **https://doi.org/10.58805/kazutb.v.2.31-1588**

A BLOCKCHAIN-ENABLED ARCHITECTURE FOR SECURE PERSONAL DATA PROCESSING AND STORAGE

[1]A.N. Tokzhigitova, [1]D.N. Yerzhembayev, [2]S.K. Zhumagulova

[1]Toraighyrov university, Pavlodar, Kazakhstan,
[2]Karaganda Buketov University, Karaganda, Kazakhstan

Corresponding author: saulesha_81@mail.ru

The article examines the problem of ensuring the integrity, confidentiality and verifiability of personal data in the context of the growth of digital services and an increase in the number of cyber incidents. The relevance of the study is due to the need to increase trust in government and corporate information systems, especially when processing biometric data. The purpose of the work is to develop and justify the architecture of a blockchain platform for the secure storage and processing of personal data based on an on-chain/off-chain hybrid model. The research methodology includes a systematic analysis of data protection requirements, the design of an information system architecture, the use of cryptographic methods (SHA-256, AES-256) and experimental verification of the implemented prototype. The proposed architecture assumes the storage of personal data in an internal secure storage (off-chain), whereas only hash values and service metadata are recorded in the blockchain journal (on-chain), ensuring the immutability and traceability of operations. As part of the work, a prototype system was implemented with access rights differentiation, mechanisms for single-time fixation of critical attributes (IIN and biometric document), control of block immutability and a procedure for detecting data substitution attempts. The experimental verification confirmed the platform' s ability to detect unauthorized changes both in files and in the block log, while maintaining the ability to adjust non-critical fields. The results demonstrate that the hybrid blockchain architecture provides a balance between performance, scalability and a high level of trust in personal data processing and can be used to modernize information systems of government and corporate structures.

Keywords: blockchain technologies, personal data, distributed registries, modern data storage systems, information security, data immutability, cryptographic mechanisms.

ЖЕКЕ ДЕРЕКТЕРДІ ҚАУІПСІЗ ӨҢДЕУ МЕН САҚТАУДЫ ҚАМТАМАСЫЗ ЕТЕТІН БЛОКЧЕЙН ТЕХНОЛОГИЯСЫНА НЕГІЗДЕЛГЕН АРХИТЕКТУРА

[1]А.Н. Токжигитова, [1]Д.Н. Ержембаев, [2]С.К. Жумагулова

[1]Торайғыров университеті, Павлодар, Қазақстан,
[2]Академик Е.А. Бөкетов атындағы Қарағанды Ұлттық Зерттеу университеті, Қарағанды, Қазақстан,
e-mail:saulesha_81@mail.ru

Мақалада цифрлық қызметтердің өсуі және киберқауіптердің көбеюі жағдайында дербес деректердің тұтастығын, құпиялылығын және тексерілуін қамтамасыз ету мәселесі қарастырылады. Зерттеудің өзектілігі мемлекеттік және корпоративтік ақпараттық жүйелерге, әсіресе биометриялық деректерді өңдеу кезінде сенімділікті арттыру қажеттілігімен байланысты. Мақсаты – тізбектегі/тізбектен тыс гибридті модель негізінде жеке деректерді қауіпсіз сақтау және өңдеу үшін блокчейн платформасының архитектурасын әзірлеу және негіздеу. Зерттеу әдістемесі деректерді қорғау талаптарын жүйелі талдауды, ақпараттық жүйенің архитектурасын жобалауды, криптографиялық әдістерді (SHA-256, AES-256) қолдануды және енгізілген прототипті эксперименттік тексеруді қамтиды. Ұсынылған архитектура жеке деректерді ішкі қауіпсіз жадта (тізбектен тыс) сақтауды көздейді, ал блокчейн журналында (тізбектегі) тек хэш мәндері мен қызмет метадеректері жазылады, операциялардың өзгермейтіндігі мен қадағалануын қамтамасыз етеді. Жұмыс шеңберінде

қол жеткізу құқықтарын саралаумен прототиптік жүйе, маңызды атрибуттарды (ЖСН және биометриялық құжат) бір реттік бекіту тетіктері, блоктың өзгермейтіндігін бақылау және деректерді алмастыру әрекеттерін анықтау тәртібі енгізілді. Эксперименттік тексеру платформаның маңызды емес өрістерді реттеу мүмкіндігін сақтай отырып, файлдарда да, блок журналында да рұқсат етілмеген өзгерістерді анықтау мүмкіндігін растады. Нәтижелер гибридті блокчейн архитектурасы өнімділік, ауқымдылық және жеке деректерді өңдеуге деген сенімнің жоғары деңгейі арасындағы тепе-теңдікті қамтамасыз ететінін және мемлекеттік және корпоративтік құрылымдардың ақпараттық жүйелерін жаңарту үшін пайдаланылуы мүмкін екенін көрсетеді.

Түйін сөздер: блокчейн технологиялары, дербес деректер, таратылған тізілімдер, деректерді сақтаудың заманауи жүйелері, ақпараттық қауіпсіздік, деректердің өзгермейтіндігі, криптографиялық механизмдер.

АРХИТЕКТУРА НА ОСНОВЕ БЛОКЧЕЙНА ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОЙ ОБРАБОТКИ И ХРАНЕНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

[1]А.Н. Токжигитова, [1]Д.Н. Ержембаев, [2]С.К. Жумагулова✉

[1]Торайгыров университет, Павлодар, Казахстан,
[2]Карагандинский Национальный Исследовательский университет имени академика Е.А. Букетова, Караганда, Казахстан,
e-mail:saulesha_81@mail.ru

В статье рассматривается проблема обеспечения целостности, конфиденциальности и проверяемости персональных данных в контексте роста цифровых сервисов и увеличения числа киберинцидентов. Актуальность исследования обусловлена необходимостью повышения доверия к государственным и корпоративным информационным системам, особенно при обработке биометрических данных. Целью работы является разработка и обоснование архитектуры блокчейн-платформы для безопасного хранения и обработки персональных данных на основе гибридной модели on-chain/off-chain. Методология исследования включает в себя систематический анализ требований к защите данных, проектирование архитектуры информационной системы, использование криптографических методов (SHA-256, AES-256) и экспериментальную проверку реализованного прототипа. Предлагаемая архитектура предполагает хранение персональных данных во внутреннем защищенном хранилище (off-chain), тогда как в блокчейн-журнал (on-chain) записываются только хэш-значения и сервисные метаданные, что обеспечивает неизменность и прослеживаемость операций. В рамках работы был реализован прототип системы с разграничением прав доступа, механизмами однократной фиксации критических атрибутов (ИИН и биометрический документ), контролем неизменности блоков и процедурой обнаружения попыток подмены данных. Экспериментальная проверка подтвердила способность платформы обнаруживать несанкционированные изменения как в файлах, так и в журнале блокировок, сохраняя при этом возможность корректировки некритичных полей. Результаты демонстрируют, что гибридная блокчейн-архитектура обеспечивает баланс между производительностью, масштабируемостью и высоким уровнем доверия к обработке персональных данных и может быть использована для модернизации информационных систем государственных и корпоративных структур.

Ключевые слова: блокчейн-технологии, персональные данные, распределенные реестры, современные системы хранения данных, информационная безопасность, неизменность данных, криптографические механизмы.

Introduction. With the development of information technology, the volume of generated and stored data is rapidly growing. This data must not only be protected, but also organized so that its integrity and manageability are ensured. As the IT infrastructure becomes more complex and

the number of users increases, the requirements for security, reliability, and data availability are becoming more restrictive. Traditional approaches to information storage and management do not always guarantee the necessary level of trust in data, which encourages the search for architectures capable of ensuring transparency, verifiability and data protection.

The relevance of the research is due to the fact that as a result of digitalization of almost all spheres of life, personal data has become one of the key types of resources.

In practice, the main threats to data are usually grouped into three key areas. The first group is illegal access and information leakage, when confidential information becomes known to persons who do not have the appropriate authority. The second is the hidden substitution or distortion of data, in which records in the system are changed without legitimate grounds. The third is a violation of accessibility, when legitimate users are temporarily or permanently deprived of the opportunity to work with data due to failures, attacks, or deliberate actions [1].

According to the National Computer Incident Response Service KZ-CERT, in 2023, about 34.5 thousand cybersecurity incidents were recorded in Kazakhstan, which is about twice as much as in 2022 and corresponds to about 17 thousand incidents the year earlier. In 2024, their number increased to 68.1 thousand, that is, almost twice as much as in 2023. According to KZ-CERT, about 34.7 thousand cyber incidents have already been registered in the first half of 2025 alone, which is almost comparable to the total for the whole of 2023. Based on these dynamics, it can be expected that by the end of 2025, the total number of incidents may exceed 70 thousand cases [2].

Organizational and technical measures are used to reduce the risk of leakage: access rights differentiation, user authentication, encryption of communication channels and stored information, control of personnel actions, leak prevention systems (DLP) and other means of protection. Backups, fault-tolerant clusters, distributed storage, duplication of critical services, and proven incident recovery procedures are used to ensure accessibility [3].

Threats of data substitution and distortion are more difficult to deal with. Unlike an obvious leak or denial of service, hidden modification of records can go unnoticed for a long time: the system continues to work, but some of the information no longer corresponds to reality. The facts of data substitution are especially critical for systems that work with sensitive information, such as biometric personal data, for example, fingerprints. In such conditions, an architecture is required that allows reliable recording of the fact and sequence of operations with data, as well as the possibility of subsequent verification of their immutability.

One of the most promising areas for solving these problems is the use of blockchain technology. Blockchain as a cryptographically secure platform provides immutability and verifiability of records, distributed data storage and the absence of a single control center [4].

The purpose of the research is to develop and substantiate the architecture of a blockchain platform for the secure storage and processing of personal data, ensuring a high level of confidentiality, integrity, and accessibility of information, as well as transparency and verifiability of data transactions. To achieve this goal, the following tasks are being solved:

- analysis of existing methods and architectures for ensuring the security of personal data using blockchain technologies;

- development of an architectural model of a blockchain platform based on a hybrid approach (on-chain/off-chain);

- implementation of a prototype system and conducting an experimental evaluation of its effectiveness in terms of safety, integrity and performance.

Materials and methods. The research methodology is based on a comprehensive interdisciplinary approach combining methods of system analysis, principles of designing modern information systems, cryptographic data protection mechanisms and experimental methods

for evaluating the effectiveness of software solutions [5]. This approach allows us to consider the developed platform not only as a software implementation, but also as an object of scientific analysis, within which architectural solutions are justified, personal data processing processes are formalized, and a quantitative assessment of the results of implementation is carried out. The use of system analysis provides identification of key requirements for security, availability and manageability of data, and also allows you to determine the relationship between the components of the information infrastructure and potential sources of threats [6].

Within the framework of the research, special attention is paid to designing the architecture of an information system focused on increasing the level of trust in data. The architectural approach is based on the principles of modularity, scalability and fault tolerance, which makes it possible to integrate blockchain components into the existing server infrastructure without significantly reducing performance. The development takes into account the requirements for the protection of personal data, including the need to maintain confidentiality and minimize the risks of information leakage. Cryptographic methods are considered as a basic tool for ensuring data integrity and authenticity, including the use of hash functions, timestamps, and access control mechanisms [5].

A key element of the methodology is the use of a hybrid data storage architecture, which involves the separation of information processing logic between internal secure storage and a distributed registry. Personal data is placed in an internal database (off-chain), which operates in a trusted environment and ensures a high level of confidentiality and manageability of information. The blockchain registry (on-chain) records exclusively cryptographic hash values, transaction identifiers, and service metadata that allows you to track the status of records and confirm their immutability. This approach avoids placing sensitive information in a distributed network, while maintaining the benefits of immutability and transparency of operations.

Using an off-chain/on-chain storage model provides a balance between system performance and the level of trust in data. The centralized repository (off-chain) is responsible for the operational processing and management of personal data, while the blockchain registry (on-chain) acts as an independent verification mechanism, recording the sequence of operations and eliminating the possibility of hidden modification of records. The research methodology involves the development of an architectural model of the platform, its software implementation and subsequent experimental verification in conditions of simulation of various scenarios of the system, including legitimate user operations and attempts to unauthorized data modification.

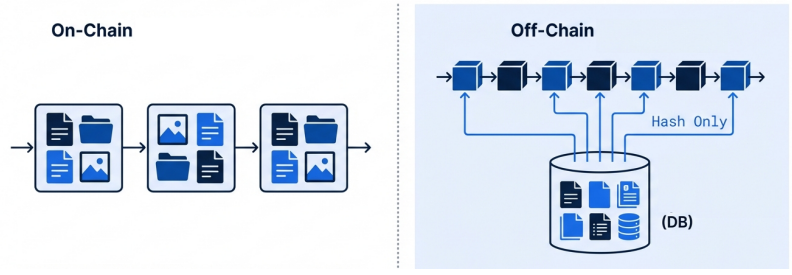


Fig.1 - Comparison of on-chain and off-chain data storage architectures

The presented Figure 1 illustrates a comparative analysis of two architectural approaches to organizing data storage in systems using blockchain technology: a fully distributed model (on-chain) and a hybrid architecture (off-chain/on-chain). As part of the traditional on-chain approach, data is placed directly in the blockchain registry, which leads to an increase in the volume of the blockchain,

complicating scaling processes and increasing demands on computing resources. In addition, hosting data on a distributed network can create additional privacy risks, especially when dealing with personal or sensitive information.

In contrast, the hybrid architecture (off-chain) involves storing the bulk of the data in an external storage or database, whereas the blockchain (on-chain) only records cryptographic hash values and metadata. This approach makes it possible to use the blockchain as a mechanism for fixing the state of data and controlling its immutability, while maintaining high system performance and manageability. Integrity verification is performed by comparing the current hash value of the data with the reference value recorded in the blockchain log, which makes it possible to detect unauthorized changes.

Thus, the proposed architectural model provides a balance between the requirements of security and scalability of the information system. Storing data in an external secure storage (off-chain) allows you to maintain confidentiality and reduce the load on the distributed registry, while using a blockchain journal (on-chain) ensures transparency of operations, cryptographic verifiability and resistance to hidden modification of information.

In a relatively short historical period, blockchain technology has gone from an experimental platform for cryptocurrencies to one of the key end-to-end digital technologies underlying modern accounting, authentication and data storage systems. If at an early stage interest in blockchain was mainly focused on electronic money, today it is considered as a universal mechanism for building reliable registries, where trust in data, verifiability of transaction history and resistance to unauthorized changes are critical [7].

This is especially true for systems working with personal data, where any hidden substitution of records can lead to serious legal consequences for citizens and organizations.

The classic definition of blockchain is associated with the work of S. Nakamoto ”Bitcoin: A Peer-to-Peer Electronic Cash System”, which proposed a decentralized payment system without a single trust center. In this system, all operations were recorded in a sequence of blocks forming a so-called ”chain”. Each block contains a hash of the previous block and a set of confirmed transactions. Over time, the idea of storing data in the form of an immutable chain of blocks linked by cryptographic hashes was transferred far beyond the financial sector and began to be considered as a general principle of organizing trusted registries [8].

From an engineering point of view, the blockchain can be represented as a distributed journal in which records are grouped into blocks, and each subsequent block contains a cryptographic fingerprint of the previous block. This seemingly simple idea ensures that any attempt to secretly change an already committed record leads to a change in the hash of the corresponding block and automatically violates the integrity of the entire subsequent chain.

Cryptographic hash functions play a key role in ensuring the immutability of information in the blockchain. The hash function accepts an arbitrary message as input and calculates a fixed-length hash value for it. An important property of the cryptographic hash function is the inability to recover the original message from the hash in a reasonable time. A small change in the input data leads to a radical change in the hash value [8].

To illustrate how the hash function works, consider a simplified example. Let’ s assume that for an arbitrary data set, we first represent it as a sequence of digits, then add up all the digits and take the result modulo 97. So, for the number 123456, the sum of the digits is 1+2+3+4+5+6 = 21, and the hash value will be 21 mod 97 = 21. If you change the source data, for example, take 78039 (7+8+0+3+9 = 27), then the hash will change: 27 mod 97 = 27. The set of possible input values is practically unlimited, while the range of hash values is only 97 options (from 0 to 96), so different input data inevitably sometimes gives the same hash. Knowing only the hash value, it is impossible to unambiguously restore the original message. There are dozens and hundreds of different combinations of numbers that lead to the same result.

Real cryptographic hash functions, such as SHA 256, use a much more complex algorithm, and the hash length is tens of bytes. This leads to the fact that in practice it is impossible to find another block of data with the same hash value or restore the original data using only one hash. From this point of view, the hash function can be considered as a one-way conversion. It is easy to calculate a hash from the data, but it is almost impossible to perform the reverse transformation.

Structurally, each block in the classic blockchain can be roughly divided into a header and a body. The header contains service fields such as the hash of the previous block, the timestamp of the block formation, the block number, the parameters of the consensus algorithm, and other metadata. The block body contains a set of records that the system places in the registry at this step. The hash of the block itself is calculated based on the block header and contents. As a result, the block can be represented as a data container with dynamic contents that are cryptographically linked to both the previous block and its own hash value [8].

The block chain is formed sequentially. First, a so-called genesis block is created, in which the ”hash of the previous block” field takes on a certain value. Each subsequent block includes the hash of the previous block in the header and thus subscribes to the entire history recorded earlier.

If an attacker wants to imperceptibly change one of the old records, he will have to change the body of the corresponding block, which will change its hash value. But this hash is already written into the header of the next block, so you will have to recalculate the next block, and so on down to the last one. Schematically, the structure of the blockchain chain is shown in Figure 2.

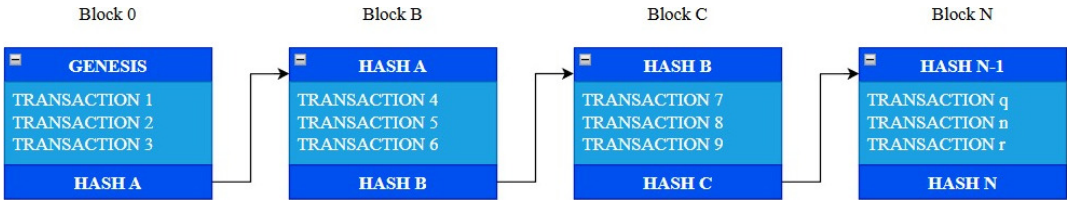


Fig. 2 - Structure of a Blockchain Chain

Having formulated general architectural approaches and chosen a private blockchain system as the target model, let’ s turn to the analysis of already implemented solutions. Practice shows that it is on the example of specific projects at the state and corporate level that the strengths and weaknesses of various models of data storage and processing organization are best demonstrated.

One of the most well-known examples of blockchain integration into public infrastructure is the use of KSI Blockchain technology in Estonia [9]. This technology protects the integrity of key government registries and electronic services. The distributed log records the hash values of records, data operations, and service metadata. At the same time, personal data of citizens, medical records and information about business activities remain in the internal systems of departments, and the blockchain is used as a means of detecting attempts to imperceptibly modify such records.

Similar ideas have been implemented in a number of projects in neighboring countries. In Georgia, the system of registration of real estate rights has been upgraded based on blockchain technologies. Confirmations of transactions with records are recorded in the blockchain, which increases citizens’ trust in the registry and reduces the possibility of imperceptible substitution of information about the owners [10]. In the Russian Federation, the Masterchain network, created by the Fintech Association with the participation of the Bank of Russia, is used as an industry platform for the financial sector. It is focused on the exchange of legally relevant messages and documents between market participants and provides an immutable transaction log [11].

In the corporate sector, blockchain is actively used to ensure the transparency of supply chains and to confirm the origin of goods. An example is the IBM Food Trust platform, where food industry participants record key product lifecycle events on the blockchain. This simplifies the audit and investigation of incidents related to product quality and safety [8]. Similar principles are used in the Everledger system, focused on accounting for the origin and movement of diamonds and other valuable assets, where the blockchain helps reduce the risk of object substitution and falsification of accompanying documentation [10].

In the healthcare sector, blockchain is used to record the fact of making changes to medical records and the history of accessing them, as well as to organize secure data exchange between medical institutions [9]. In the field of education, projects are being implemented to use blockchain to store academic achievements and issue digital certificates and diplomas.

For Kazakhstan, blockchain is gradually becoming one of the tools for developing e-government and increasing trust in digital services. Separate pilot projects are aimed at using a distributed registry to record legally significant actions in government information systems, log data accesses, and monitor the integrity of registry entries. Solutions based on NFT diplomas are being implemented to verify the authenticity of educational documents through a distributed registry [12].

These examples show that in most practical solutions, blockchain acts as a superstructure over traditional information systems responsible for the immutability and the traceability of critical records. The most stable schemes are those in which sensitive data continues to be stored in domain repositories, and their hashes and transaction logs are transferred to the distributed registry. This approach seems to be the most justified when developing a platform for working with personal data.

For the tasks solved in this research, an important conclusion is that blockchain technology has shown its viability as a means of ensuring the immutability and traceability of data transactions in real systems, including scenarios with the processing of personal information.

Within the framework of the blockchain platform being developed, the latter is considered as a trusted transaction log, which records the facts of registration, updating and use of fingerprint drawings, while the images themselves are stored in external repositories. Schematically, the proposed hybrid architecture of the platform is shown in Figure 3.

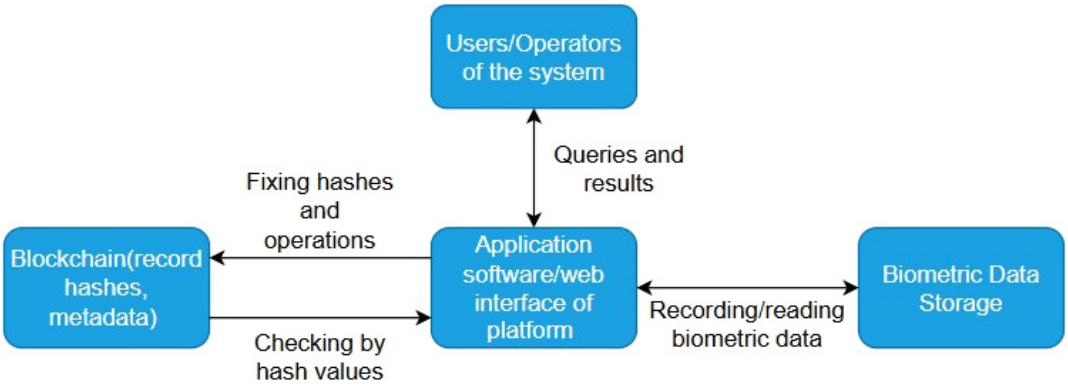


Fig.3 - Hybrid platform architecture. Blockchain magazine and secure Biometric data storage

To evaluate the effectiveness of the developed blockchain platform, a set of quantitative and qualitative criteria was introduced. The evaluation methodology includes analysis of integrity verification accuracy, transaction processing performance, scalability under increasing load,

and resistance to unauthorized modification of records. In addition, threat modeling was performed based on confidentiality, integrity, and availability principles.

Table 1 - Evaluation criteria and performance metrics of the blockchain platform

Evaluation category	Metric
Integrity	hash mismatch detection rate
Performance	average transaction processing time
Scalability	blockchain growth rate
Security	unauthorized modification detection
Availability	recovery time after failure

Results and discussion. In international practice, personal data refers to any information related to an identified or identifiable individual: last name, first name, identifiers, location information, online identifiers, health data, and other characteristics that directly or indirectly identify a person. This approach is enshrined, in particular, in the EU General Data Protection Regulation (GDPR, Regulation (EU) 2016/679 [13].

GDPR is considered one of the most well-developed and stringent regulations in the field of personal data protection. It sets out a set of basic principles for data processing, which in fact serve as a guideline for many other jurisdictions. These are legality, integrity and transparency of processing, limitation of purposes, minimization of data, accuracy, limitation of retention periods, integrity and confidentiality, accountability of the data storage operator.

A set of rights of the data subject, enshrined in European law, is also essential: the right to access their data, to correct, delete, restrict processing, data portability, as well as the right to object to processing. These rights are combined with the obligations of operators to obtain unambiguous and informed consent, document processing processes, and notify supervisors and subjects of security breaches and unauthorized access. In addition to GDPR, other regulatory models are actively developing in the world. For example, the Council of Europe Convention 108+ on the Protection of Persons in connection with Automated Processing of Personal Data [14].

Technological requirements for the architecture of information systems are becoming a separate area of development. The concepts of minimizing the attributes collected, data encryption, separation of identifiers and profile content, as well as mandatory maintenance of detailed access and change logs are being implemented everywhere. These approaches are closely related to architectural solutions, including the use of distributed registries and blockchain technologies, which allow for provable audit of operations while maintaining controlled access to the data itself.

In the Republic of Kazakhstan, the basic act in the field of personal data is the Law ”On Personal Data and their Protection” dated May 21, 2013 No.94-V. The purpose is to ensure the protection of human and civil rights and freedoms in the collection and processing of personal data. The law introduces the concepts of personal data, the subject of personal data, the owner, the operator and a third party, thereby formalizing the roles of the participants in the processing. Personal data is defined as information related to a specific or identifiable data subject, recorded on electronic, paper or other material media [15].

The law establishes the rights of a personal data subject, such as the right to receive information about the nature and scope of the information being processed, to access their data, to clarify, block, destroy it in case of illegal processing or the need to bring it into line with the actual state of affairs.

At the same time, the responsibilities of the owners and operators of personal data are established to ensure confidentiality, take legal, organizational and technical protection measures, restrict access to authorized persons and keep records of processing operations.

An important role is also played by the Law of the Republic of Kazakhstan ”On Informatization”, which establishes requirements for the construction and operation of state information systems, for the deployment of information and communication infrastructure, as well as for ensuring information security. On its basis, requirements are introduced for the localization of certain categories of personal data, the deployment of critical systems on the territory of Kazakhstan and the use of certified protective equipment [16].

In addition, the cybersecurity Concept ”Cyber Shield of Kazakhstan”, approved by a government decree, defines strategic guidelines in terms of data security. The document defines the goals for the protection of electronic information resources, state information systems and communication networks, and also establishes the need to develop a national cybersecurity infrastructure [17].

In the context of digitalization, Kazakhstan is also implementing the Digital Kazakhstan state program, which focuses on expanding electronic services, developing e-government and improving the level of digital literacy of the population. All this leads to a steady increase in the volume of personal data processed and the complexity of their storage infrastructure [18].

Despite the existence of basic legislative regulation and strategic documents, the practice of recent years shows that serious problems remain in the field of personal data protection. The rapid growth of electronic services and the integration of government information systems are outpacing the development of information security mechanisms and law enforcement practices, and the level of awareness of citizens about the risks remains insufficient. At the same time, there are not only threats of unauthorized disclosure of information, but also risks of deliberate substitution or distortion of data in key registries.

Major incidents involving the leakage of personal data of a significant part of the population of Kazakhstan have attracted particular public attention. Arrays of data containing information about millions of citizens appeared in the public domain. It follows from official statements and media publications that leaks are often associated not so much with external hacker attacks as with the unauthorized use of legitimate accounts with access to government databases. A separate group consists of incidents involving the deliberate entry of deliberately false information into government information systems. In a number of high-profile cases, employees of public service centers, using official access, changed or added information to registers, which allowed them to illegally process documents, register property, or receive services on behalf of others. Such actions lead not only to the violation of the rights of specific citizens, but also undermine the credibility of state registers as a whole [19,20].

With the dominance of a centralized architecture based on traditional databases and broad administrator authority, any transaction logs, access protocols, and even primary records can potentially be modified by administrators themselves or by intruders who have obtained their authority. This makes it more difficult to investigate incidents, reduces confidence in the results of electronic procedures, and makes it more difficult to prove the facts of unauthorized interference, including data substitution.

The listed trends and problems directly correlate with the tasks set in this master’ s project. On the one hand, requirements have already been formulated at the state level to ensure the confidentiality, integrity and accessibility of personal data, to localize critical information resources and increase the transparency of processing processes. On the other hand, the practice of massive leaks and internal abuse of access demonstrates that formal requirements and traditional architectures alone are not enough to ensure an adequate level of trust in the personal data storage infrastructure.

The development of a blockchain platform for storing and processing personal data is considered

by the author as one of the options for responding to these challenges. It should combine the mechanisms of cryptographic protection and immutability of records with the possibility of flexible access control and integration.

However, to justify the relevance of developing and implementing a blockchain platform for processing and storing personal data, it requires not only an analysis of internal technical and legal parameters, but also a systematic assessment of the environmental factors in which the proposed solution will function.

One of the most common tools for this kind of assessment is PEST analysis, which allows a systematic description of the political, economic, social and technological factors affecting the implementation and operation of a platform for working with biometric data. The results of the analysis are presented in Table 2.

Table 2 - PEST analysis of the external environment

Factor	Description of the influencing factors	Influencing the project
Political	The state digitalization programs of the Republic of Kazakhstan, the development of e-government and public services in online format stimulate the implementation of solutions for the secure processing of personal data. At the same time, the requirements for the protection of personal data and critical information and communication infrastructure are being strengthened.	Prerequisites are being created for the testing of the blockchain platform in government and quasi-government organizations, but expectations for compliance with regulatory requirements and information security standards are increasing.
Economic	The increase in direct and indirect losses from cyber incidents makes investments in personal data protection economically justified. At the same time, the budgets of many organizations, especially regional and educational ones, are limited, which complicates the implementation of complex and expensive solutions.	The developed platform can be positioned as a relatively inexpensive and step-by-step solution, but it is necessary to justify economic efficiency and low total cost of ownership.
Social	The population is increasingly using digital services and expects organizations to take a responsible approach to personal data protection. At the same time, distrust of centralized repositories of biometric information persists, and society's sensitivity to incidents of data leakage and substitution increases.	The transparency and verifiability of data transactions provided by the blockchain platform can increase the trust of users and stakeholders. At the same time, work is required to improve the culture of information security and explain the principles of the system.
Technological	The development of open blockchain platforms, cryptographic libraries, and information security tools simplifies the creation of prototypes and pilot solutions. At the same time, the computing power of potential violators is growing, and new threats are emerging (including those related to the development of quantum computing).	The architecture of the platform should be flexible, with the possibility of updating cryptographic algorithms and scaling to a distributed network. At the project stage, it is justified to use local solutions with minimal requirements for communication channels and infrastructure.

Thus, the results of the PEST analysis show that the external environment is generally favorable for the development and implementation of a blockchain platform for working with biometric personal data. Digital transformation and government programs are stimulating the demand for transparent and reliable solutions, users are increasingly sensitive to issues of personal

information protection, and the technological base is gradually being formed. At the same time, there are still a number of limitations related to budgetary and infrastructural resources, a shortage of qualified specialists and the need for constant updating of the applied information security tools. These factors must be taken into account when choosing the architecture of the solution and planning the stages of its implementation.

The main functional idea of the solution is to combine classical data storage mechanisms with a local blockchain log, which records control values for critical fields and allows detecting substitution attempts. The system implements the differentiation of data into two groups:

1) critical data subject to fixation and integrity control through the block chain;

2) mutable attributes that can be adjusted by the user and do not affect the reference value.

The critical data includes the individual identification number (IIN) and the fingerprint image (hereinafter referred to as the document). A dataHash reference value is generated for this data. The remaining fields of the card (full name, employee's photo) can be changed without recalculating the reference value and without changing previously fixed blocks.

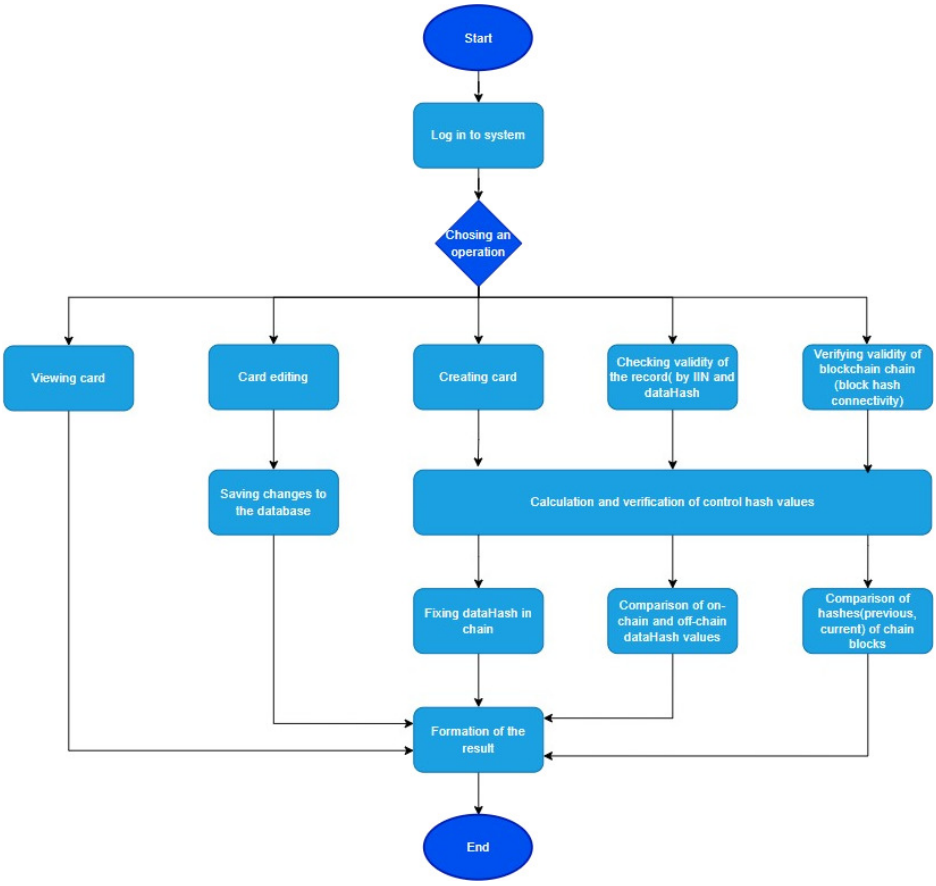


Fig.4 - Generalized flowchart of the program operation algorithm

In the implementation under consideration, the deployment outline includes a server application (NestJS/Node.js), MySQL database management system (users, employees, blockchain tables) and /public/uploads file storage for uploading files (photos and document). The key design decision is that metadata and control values are stored in the database, and the immutability of critical parameters is ensured by fixing the calculated dataHash in the blockchain journal.

At the application launch stage, the BlockchainService service loads the chain from the blockchain table, and in the absence of data, generates a genesis block. To ensure reproducibility of hash calculations during repeated launches, stable JSON serialization (key sorting) is used, which eliminates the dependence of the result on the internal order of fields in MySQL and allows validating the chain correctly after restarting the application.

The critical fields of the employee' s card are recorded once when creating an entry. documentHash is calculated as SHA-256 from the contents of the document file, then dataHash is formed as SHA-256 from the result of concatenation of iin and documentHash. The resulting value is stored in employees and simultaneously written to a new block of type EMPLOYEE_DATA (EmployeeID, dataHash, meta). During subsequent editing of the card, only non-critical fields can be changed, whereas the IIN and the document file are not overwritten and do not involve changing previously fixed blocks. The interface of the procedure for adding an employee and making changes is shown in Figure 5.

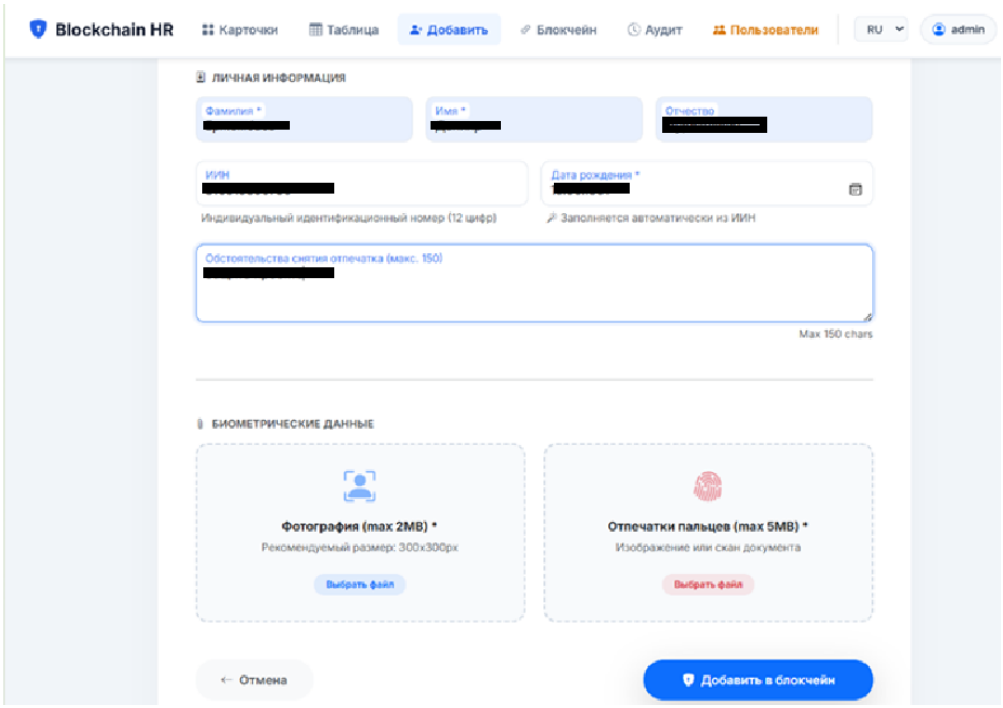


Fig.5 - Interfaces for adding a new card and making changes

A separate mechanism to reduce the risk of personal data compromise is IIN encryption when stored in the employees table. The value is encrypted in the EmployeesService using AES-256. All verification logic is performed at the application level after decryption, which allows you to maintain the verifiability of the dataHash and at the same time not store the IIN in the clear in the database.

The scientific novelty of the research lies in the development of a hybrid blockchain architecture for biometric personal data protection combining off-chain confidential storage and on-chain integrity verification mechanisms.

Conclusions. test the developed blockchain platform, it is proposed to determine its ability to implement the main tasks of the project described earlier. It is advisable to consider testing as a check of the achievability of the specified criteria at the level of system behavior.

The results of the verification of the blockchain platform against the project criteria are shown in Table 3.

Table 3 - Criteria for matching tasks

Expected property (task)	Implementation in the system	Engineering sign of completion
Delimiting access to operations	users (role, permissions) + AuthMiddleware: transfer rights to templates and route protection	an attempt to access /employees/create and /audit without rights results in a denial
One-time fixation of controlled data	createEmployee: calculation of documentHash and dataHash; writing dataHash to employees and the EMPLOYEE_DATA block	The dataHash benchmark is present simultaneously in employees.dataHash and blockchain.data.dataHash
Immutability of critical card attributes	editEmployee: the existing iin and documentHash are used; the dataHash is recalculated from unchanged values	after editing, only non-critical fields are changed. The dataHash and the corresponding block remain unchanged
Detecting document file substitution	verifyEmployee/verify-all: reading a file from /public/uploads and recalculating documentHash and dataHash	A discrepancy between computedHash and onChainHash or a discrepancy between documentHash is recorded as a violation
Detecting Block log tampering	isChainValid: recalculating the hash of each block and checking the previousHash	any change to data/previousHash/hash/nonce invalidates the chain
Auditing changes to non-critical fields	employees.modificationCount and updatedAt are updated each time the allowed attributes are edited	It is possible to establish the fact and number of modifications without changing the fixed control values

Unlike traditional centralized database systems, where administrators may potentially modify logs or records, blockchain-based architectures provide cryptographic immutability and transparent auditability of operations.

Table 4 - Comparison of traditional DBMS and blockchain-based architecture

Aspect	Traditional DBMS	Proposed platform
Hidden modification detection	Limited	Automatic
Log immutability	Low	High
Audit transparency	Partial	Full

Despite the advantages of the proposed architecture, several limitations remain, including dependence on trusted off-chain storage, blockchain growth over time, and computational overhead associated with cryptographic hashing.

Thus, the evaluation of the implementation shows that the declared functions of the project are provided by specific mechanisms:

- access to operations is regulated by the user’ s role and list of permissions;

- record integrity control relies on a calculated dataHash associated with the IIN and the contents

of the document file;

- immutability of the block log is ensured by hash/previousHash connectivity and hash recalculation verification;

- changing non-critical fields does not recalculate control values and does not change previously created blocks.

References

1. Egoshin, N.S.; Konev, A.A.; Shelupanov, A.A. Model of Threats to the Integrity and Availability of Information Processed in Cyberspace//Symmetry. - 2023.-Vol.15:431. DOI 10.3390/sym15020431.

2. Incident statistics. KZ-CERT (National Computer Incident Response Service of the Republic of Kazakhstan): URL: https://cert.gov.kz/press_club/infographics. -Date of access: 06.02.2026. [in Russian]

3. Swanzy P. N., Arnold M.A., Edward D. A.2024. Data Security Framework for Protecting Data in Transit and Data at Rest in the Cloud// Current Journal of Applied Science and Technology. - 2024.-Vol.43(6). - P.61-77. DOI 10.9734/cjast/2024/v43i64387.

4. The Law of the Republic of Kazakhstan. -2015. -No.418-V SAM ”On informatization” // Adilet: information and legal system of regulatory legal acts of the Republic of Kazakhstan. URL: https://adilet.zan.kz/eng/docs/Z1500000418. -Date of access: 06.02.2026.

5. Awan M., Alam A., Khan R.A. et al. A generative AI-driven cybersecurity framework for small and medium enterprises software development: an ANN-ISM approach. - Sci Rep.- 2026.- 16(1): 9813. DOI 10.1038/s41598-026-37614-8.

6. Kioskli K., Grigoriou E., Islam S., et al. A risk and conformity assessment framework to ensure security and resilience of healthcare systems and medical supply chain// International Journal of Information Security. -2025. -Vol.24: 90. DOI 10.1007/s10207-025-01009-z.

7. Abad-Segura E., Infante-Moro A., González-Zamar M-D., López-Meneses E. Blockchain Technology for Secure Accounting Management: Research Trends Analysis//Mathematics. -2021. -Vol.9(14):1631. DOI 10.3390/math9141631.

8. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System// Cryptography Mailing list at https://metzdowd.com. -2009. URL: https://www.researchgate.net Date of access: 06.02.2026.

9. KSI blockchain in Estonia// e-Estonia. URL: https://e-estonia.com/solutions/cyber-security/ksi-blockchain/. -Date of access: 06.02.2026.

10. Alobid, M.; Abujudeh, S.; Szűcs, I. The Role of Blockchain in Revolutionizing the Agricultural Sector //Sustainability. – 2022. -Vol.14: 4313. DOI 10.3390/su14074313.

11. Zheng Z., Xie S., Dai H., Chen X., Wang H. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends // 2017 IEEE International Congress on Big Data. -2017. -P.557–564. DOI 10.1109/BigDataCongress.2017.85.

12. The First in Central Asia: Narxoz will issue NFT diplomas. -2023. URL:https://vernycapital.com/en/the-first-in-central-asia-narxoz-will-issue-nft-diplomas/. -Date of access: 06.02.2026.

13. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. -Date of access: 06.02.2026.

14. Council of Europe, “Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108+),” Council of Europe Treaty Series (CETS No.223), 2018. URL: https:// grokipedia.com/page/convention_for_the_protection _of_individuals with_regard_to_automatic_processing_of_personal_data -Date of access: 06.02.2026.

15. Republic of Kazakhstan, “Law on Personal Data and Their Protection,” No.94-V, May 21, 2013. URL: https://adilet.zan.kz/eng/docs/Z1300000094. -Date of access: 06.02.2026.

16. The Law of the Republic of Kazakhstan ”On Informatization”. -2015. -No.418-V. URL: https://adilet.zan.kz/eng/docs/Z1500000418.-Date of access: 06.02.2026.

17. Government of the Republic of Kazakhstan, “Cyber Shield of Kazakhstan: Cybersecurity Concept,” 2020. URL: https://www.gov.kz/memleket/entities/knb/activities/250?lang=en. -Date of access: 06.02.2026.

18. Government of the Republic of Kazakhstan, “Digital Kazakhstan State Program,” 2025. URL: https://www.kmge.kz/en/state_program_en/. -Date of access: 06.02.2026.

19. Delgado-Mohatar, O., Fierrez, J., Tolosana, R., Vera-Rodriguez, R. Blockchain and Biometrics: A First Look into Opportunities and Challenges. In: Prieto, J., Das, A., Ferretti, S., Pinto, A., Corchado, J. (eds) Blockchain and Applications. BLOCKCHAIN 2019. Advances in Intelligent Systems and Computing.-2019.-Vol.1010.-P.169-177.DOI 10.1007/978-3-030-23813-1_21.

20. Ghafourian, M.; Sumer, B.; Vera-Rodriguez, R.; Fierrez, J.; Tolosana, R.; Moralez, A.; Kindt, E. Blockchain and Biometrics: Survey, GDPR Analysis, and Future Directions. arXiv 2025, arXiv:2302.10883 [cs.CV]. https://arxiv.org/abs/2302.10883.

Information about the authors

Tokzhigitova A.- PhD, Associate Professor, Toraighyrov university, Pavlodar, Kazakhstan, e-mail: tokzhigitova.an@teachers.tou.edu.kz;

Yerzhembayev D. - Master of Engineering and Technology, Toraighyrov university, Pavlodar, Kazakhstan, e-mail: yerzhembayev@gmail.com;

Zhumagulova S. - PhD, Associate Professor, Karaganda Buketov University, Karaganda, Kazakhstan, e-mail: saulesha_81@mail.ru.

Сведения об авторах

Токжигитова А.Н.- PhD, ассоциированный профессор, Торайгыров университет, Павлодар, Казахстан, e-mail: tokzhigitova.an@teachers.tou.edu.kz;

Ержембаев Д.Н. магистр техники и технологии, Торайгыров университет, Павлодар, Казахстан, e-mail: yerzhembayev@gmail.com;

Жумагулова С.К.- PhD, ассоциированный профессор, Карагандинский Национальный Исследовательский университет имени академика Е.А. Букетова, Караганда, Казахстан, e-mail: saulesha_81@mail.ru