Journal of Electrical Engineering and Information Technologies, Vol. **8,** No. 2, pp. 93–100 (2023)
Article 210
In print: ISSN 2545–4250
Received: November 4, 2023
On line: ISSN 2545–4269
Accepted: December 6, 2023
UDC: 004.652.3-027.544/.545:[002:378.2
https://doi.org/10.51466/JEEIT2382210093r

Original scientific paper

DATABASE DESIGN IN A BLOCKCHAIN-BASED SYSTEM FOR GENERATING AND VERIFYING ACADEMIC CREDENTIALS

Avni Rustemi[1,2], Vladimir Atanasovski[1], Aleksandar Risteski[1]

[1]*Faculty of Electrical Engineering and Information Technologies, "Ss. Cyril and Methodius" University in Skopje, Rugjer Bošković bb, P.O.Box 574, 1001 Skopje, Republic of North Macedonia*
[2]*Department of Informatics, Faculty of Mathematics and Natural Sciences, Tetovo University, 1200 Tetovo, Republic of North Macedonia*
avni.rustemi@unite.edu.mk

A b s t r a c t: The management of data and their storage in the database for an indefinite time, whether in a centralized or decentralized form, is of great importance for the operation of the system and its efficiency. One of the challenges that decentralized systems are facing is the lack of decentralized management of databases, because they have limitations in terms of inserting, updating, and deleting data. The generation of diplomas is a more sensitive issue and one of the reasons why it has not been digitalized to date is the fact that centralized systems do not offer sufficient security for the generation of such documents, since the same systems must be controlled by third people or administrators of the system, where there is also the suspicion of data misuse. Through the paper, we will compare centralized and decentralized databases, the operation of decentralized databases, and finally present a scheme of our vision about the operation of the database in the blockchain-based system for the generation and verification of diplomas.

Key words: entralized database: decentralized database: blockchain system: diploma

ДИЗАЈН НА БАЗА НА ПОДАТОЦИ ВО СИСТЕМ ЗАСНОВАН НА БЛОКЧЕЈН ЗА ИЗДАВАЊЕ И ПРОВЕРКА НА АКАДЕМСКИ ДОКУМЕНТИ

А п с т р а к т: Управувањето со податоците и нивното складирање во базата на податоци на неопределено време, било во централизирана или децентрализирана форма, е од големо значење за функционирањето на системот и неговата ефикасност. Еден од предизвиците со кои се соочуваат децентрализираните системи е немањето децентрализирано управување со базите на податоци, бидејќи имаат ограничувања во однос на вметнување, ажурирање и бришење податоци. Издавањето дипломи е чувствително прашање и една од причините зошто до денес не е дигитализирано е фактот што централизираните системи не нудат доволна сигурност за издавање такви документи, бидејќи истите тие системи мора да бидат контролирани од трети лица или администратори на системот, бидејќи и кај нив постои сомнеж за злоупотреба на податоците. Во трудот ги споредуваме централизираните и децентрализираните бази на податоци, работата на децентрализираните бази на податоци и на крајот претставуваме шема за тоа како ја замислуваме работата на систем за издавање и верификација на дипломи базиран на блокчејн.

Клучни зборови: централизирана база на податоци; децентрализирана база на податоци; систем блокчејн; диплома

1. INTRODUCTION

Very important for a higher education institution is efficient management, real-time generation, safe and secure storage of data in certain databases for a longer period of time. During the creation of a certain system, it is very important to choose the platform where the same will be implemented, the auxiliary equipment and to provide opportunities for updating the system without affecting the data in

the database. Also, the reuse of architectures, models and certain parts of the code, speed up the creation time, reduce costs, and make the system easily adaptable to different platforms. Blockchain technology is seen as a solution for digitizing services as a whole, including the process of generating and verifying diplomas. This technology offers distributed data processing networks, generating data automatically without the intervention of the third parties. However, despite the fact that blockchain technology is finding a wide application in many fields, in education, and especially in higher education institutions, it is encountering many difficulties. From the systematic literature review [1], it turns out that despite many challenges and obstacles, one of the main ones is that there is a lack of developers' interest in blockchain programming, and the creation of data management systems. There is a lack of interoperability between decentralized databases and programming languages with which the system is developed as a whole, and the smart contract in particular. The generation must be connected with the preliminary processes in the blockchain system, from enrollment process, evaluation to obtaining grades and credits in certain subjects. Smart contract and blockchain programming have certain limitations, because they are compatible only on the Ethereum platform, while the same platform cannot achieve the maximum performance, due to many characteristics. The automation of services has its challenges and limitations, where for every mistake the data cannot be updated, and requires deletion and regeneration of the same. Blockchain systems require much more auxiliary tools to carry out transactions than centralized systems. And finally, there is a lack of architectures and concrete standard models that describe the processes of generation and verification of diplomas, respectively there is a lack of literature. Table 1 gives a description of the characteristics that blockchain-based systems possess, where, among other things, the same should offer high security, identity protection, transparency of services, their classification and even the generation and verification of academic documents quickly and safely [2].

Despite the fact that so far several educational platforms have been implemented for the educational process in higher education institutions [3], even for the verification of diplomas [4], however, most of them have been tested with several data and there is no research that shows that the same are being used or have shown good or bad results during application.

T a b l e 1

Characteristics of blockchain-based system

Characteristics	Description
Privacy	Divide the services into private and public
Traceability	The data is immutable, alerts any change
Transparency	Services are transparent, easy and fast verification of diplomas
Safety	Intelligent and cryptographic mechanisms
Trust	Gain people's trust from the characteristics it possesses

2. SMART CONTRACT USABILITY

Presented in the simplest form, the process of communication between the user and the blockchain network is given the Figure 1. Between the user interface and the execution of the smart contract is the web3 provider, as well as between blockchain network and smart contract it is blockchain api. Smart contracts are compatible with Solidity as a programming language while their execution is done on the Ethereum platform, as one of the platforms most compatible with smart contracts [5]. Each smart contract represents a digital object in a blockchain system. Every transaction on the Ethereum platform represents a smart contract [6].

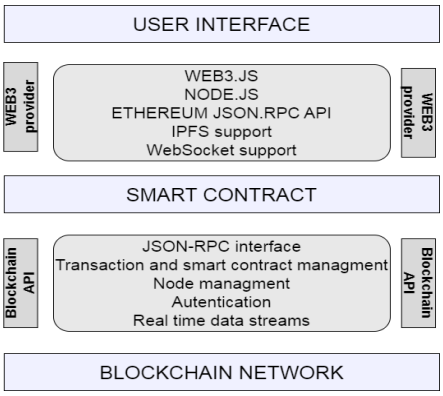

Fig. 1. Interaction between blockchain network and user

To support the work of smart contract developers, online environments called Remix IDE have been created. Smart contract can be used in banking systems for loans, automatic management of bank transactions related to large companies, in medicine for health insurance, automatic allocation of financial resources to clinics based on fulfilled conditions, in health insurance for damage assessment

based on the conditions as well as many other areas, where a fair judgment with preliminary conditions is needed [7].

In medicine for the automatic processing of doctors' requests, for the allocation of financial resources to clinics. In electronic commerce for the management of payments, online orders. In aviation for automatic flight management, ticket booking, payment system.

Regarding the application of smart contract in higher education institutions, they are used in digital certificates, for the process of storing, verification and validation of certificates, as a supporting tool for many processes such as online learning, various online trainings for students and academic staff, the management of the bookstore in the institution of higher education, student payments, students' knowledge evaluate, student evidence, registration process [8]. Among the many challenges that we have mentioned in [9], there is also the challenge of standardizing smart contracts and their automatic generation on the Ethereum platform, since for higher education institutions, such a phenomenon is limited. Any error in the data must result in its regeneration, then the cost doubles, since the cost of maintaining the blockchain system itself is much higher than the cost of centralized systems. The immutability of smart contracts makes it even more difficult to adapt and reuse them.

Figure 2 presents the smart contract execution process, starting from the invoking smart contract until the successful implementation or not, depending on the money we have in the wallet, as long as Figure 3 summarizes some of the challenges facing blockchain technology in higher education institutions (HEI).

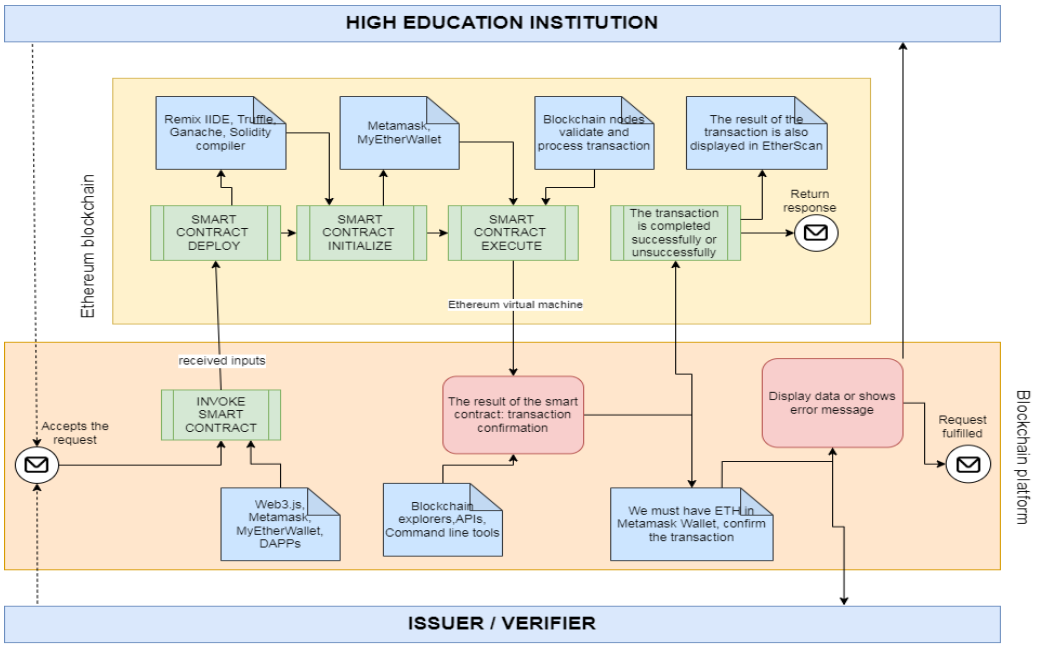

Fig. 2. Smart contract executing process

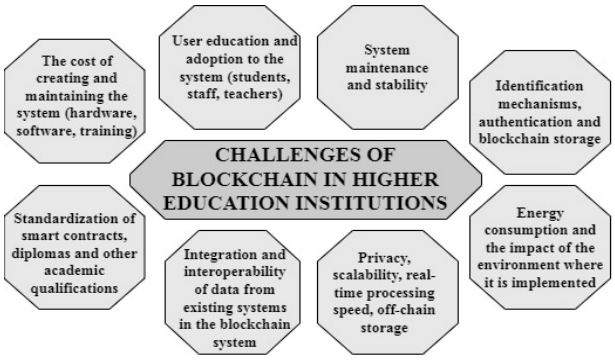

Fig. 3. Challenges of blockchain in HEI

3. DECENTRALIZED SYSTEM AND DATABASES CHARACTERISTICS

The development of information computer technologies (ICT) and the digitization of services has caused that in almost all spheres of life, most services are digitized and generated electronically through various devices. Compared to the past, with traditional learning methods, where students have had the opportunity to acquire knowledge only through books, nowadays they can access any information they need in digital form in a very short time

through digital devices [17]. Learning Management Systems (LMS) also facilitate the educational process in higher education institutions. Today, the distribution of literature, electronic learning, real-time tasks, electronic evidence, distance learning, and the classroom, are just some of the many opportunities that students have thanks to LMS. It is worth emphasizing a great help of Youtube in this direction, where through videos the teachers very easily share to the students the lessons with relevant explanations, which luckily the same can be shared into the LMS and in the classroom [20]. Centralized systems always have a third intermediary in the generation of each service, respectively every service must be generated by the administrative staff, or other responsible persons, up to the administrator. The application of artificial intelligence in centralized systems has helped to develop processes in an independent form, respectively intelligent devices have the ability to learn from human actions, however, centralized systems must have an administrator who controls every process in the system and is responsible for whatever happens in the system. [18]. Despite the fact that there are many developments in centralized systems, offering fast, secure data processing, by the word itself we mean that someone has access to the data, and he is definitely the system administrator. Precisely for this reason, privacy, authentication and even the possible misuse of data and the use of the same for malicious purposes are put into question. What is more worrying, in such systems, in institutions of higher education, in addition to the administrator, IT managers and the higher management of the institution have access, which means that the possibility of misuse is even higher. And the system lacks a traceability that cannot be hidden, but that shows precisely every possible change of the data and by whom it was made [19].

Decentralized systems in comparison to centralized ones have some advantages, however, the respective limitations. In the following, we will mention some differences of the same to continue with the types and clarification of decentralized databases.

- Privacy, transparency, traceability, identity [10]. Lack of self-identification mechanisms, the target of numerous attacks, the cause of transparency
- Reuse and adaptability, decentralization and distributed ledger [11].Reuse depends on standardization, adaptation is difficult especially from centralized to decentralized systems, decentralization has more deficiencies in storage for a longer period of time and response in real time.
- Immutability of smart contract and blockchain platform [12, 14]. Every error requires the regeneration of the smart contract, smart contracts are only compatible with the Ethereum platform, which limits it in many features
- Safety and cost, electricity consumption, maintenance [13, 14]. The generation and verification of diplomas have a low cost, but their maintenance for a longer time is expensive, because blockchain itself is an expensive technology.
- Decentralized data storage, off-chain storage [15]. All current systems work with centralized databases, and adaptation to decentralized databases is very difficult.
- User interface, compatibility, guide and documentation [15]. The design of the system must be compatible for all intelligent devices, lack of documentation and experts in this field.
- Fast generation, verification, online support all the time [13–15]. It requires high costs, because it is in direct proportion to the energy consumed and the blockchain auxiliary equipment used for the system. There is a lack of programmers and experts in this direction

As for decentralized databases, there are several of them, whether as file storage, cloud platforms or even blockchain databases, among which we will mention:

- Interplanetary file storage (IPFS), which has shown good results in data management in certain capacities with blockchain systems, which is based on the peer-to-peer protocol, and uses the addressing method to store and receive data.
- BigChainDB, combines blockchain technology with decentralized databases. In fact, it is used for storing and processing the largest amounts of data in the blockchain network.
- Swarm, is a peer to peer protocol, which stores data in an unchangeable form. It runs on the Ethereum platform [16].
- Cassandra, is a NoSQL database, which is based on the addition of more nodes for data storage, which is actually known as the continuation of SQL databases, and those who are familiar with it, find it very easy to use Cassandra.
- ChainifyDB, is a blockchain layer that has recently been used as an auxiliary tool for existing blockchain databases, and is also used as a connecting tool for data in the blockchain network.
- CovenantSQL, is a blockchain database that attempts to combine the characteristics of central-

ized databases with the characteristics of blockchain, giving primary importance to the issue of privacy and data security.

- Modex Blockchain Database (BCDB), developed by the company Modex, known for its research in the field of blockchain. It aims to combine the advantages of traditional centralized databases with the immutability and transparency of the blockchain.
- Postchain, developed by ChromaWay company, specialized in smart contract and decentralized applications. It also aims to combine relational databases with blockchain technology.

4. DATABASE DESIGN FOR BLOCKCHAIN SYSTEM

Very important for the selection of the blockchain database are several factors such as the data structure, consensus mechanisms, node types, block structure, determining the amount of data to be stored and processed, up to security strategies, scalability, transparency. In order to have a clearer overview of the functioning of the blockchain system, for the generation and verification of diplomas, including the database, we will present a class diagram through Figure 4, as part of the UML diagrams. In this diagram, we have simultaneously presented the most important elements that are necessary for the implementation of the blockchain system for the generation and verification of diplomas, including the blockchain database. Both processes, the generation of diplomas and the verification, are connected to the blockchain database. And the blockchain database works through the Ethereum blockchain platform, with which the data is mined, validated, encrypted and sent in a distributed form to the respective database. Very important for the efficient generation of diplomas, is the insertion of data in a detailed form in the system, taking special care for all information not to be mistaken, because every mistake has its own cost in the blockchain system. It is of little importance whether Homomorphic Encryption, Zero-Knowledge Proofs or Elliptic Curve Cryptography will be used, as they aim to preserve data security, privacy, collaboration and database management in a protected way [16].

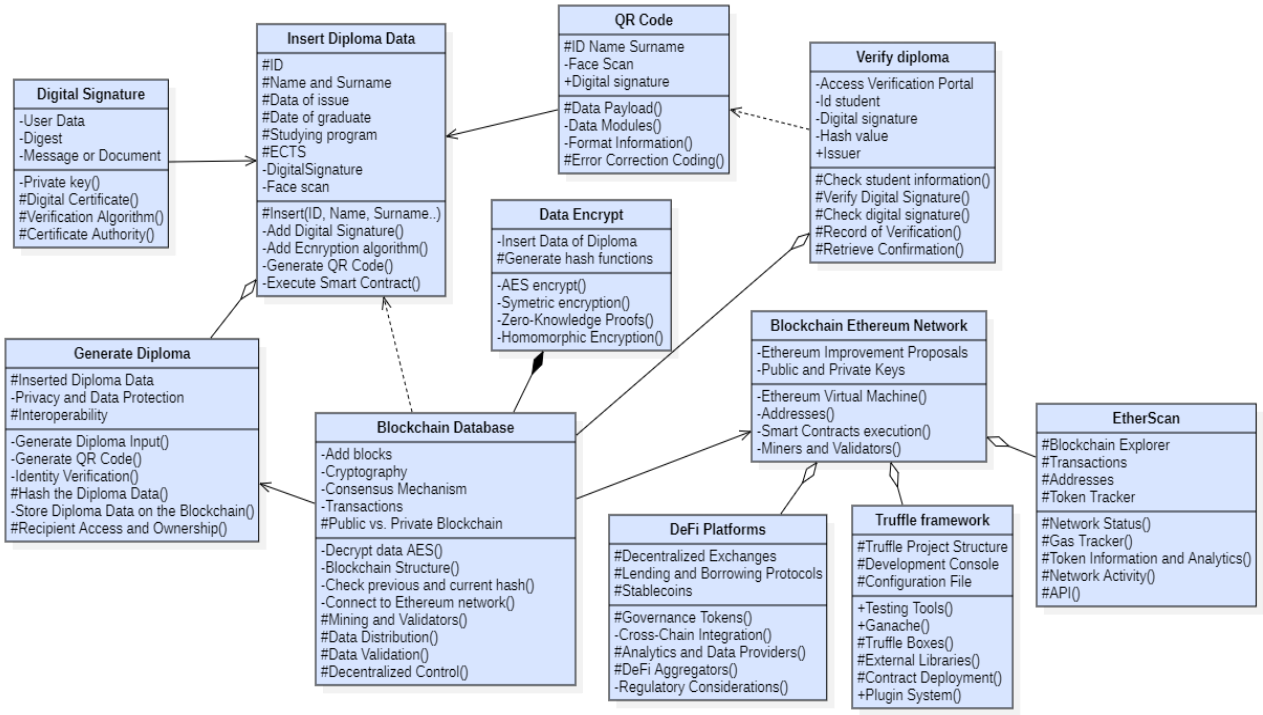

Fig. 4. Class diagram for blockchain system and database design

The digital signature is very important during the insertion of the data in the database, as the same must be signed at the end by the student but also by the institution's management. As part of the digital signature, we have operations such as: private key generation, digital certificates, certificate authorization, algorithms for verification. It is very important for quick verification to attach the QR code to the

generation of the diploma. QR code, on the one hand, verifies the immutability of the data, since all the data is generated in the QR code, but on the other hand, it is very easy to verify the data, since the data that is part of the QR code can be verified through very easy applications. Within the blockchain database, we have presented the following operations: data decryption and encryption, blockchain structure for data storage and distribution, checking previous and current hash values, connection to Ethereum blockchain network, mining and validation, decentralized data control. Additionally to description of the class diagram for blockchain system and database design presented on Figure 4, Figure 5 illustrates the sequential diagram of the process of generating the diploma in electronic form, digitally signed by both the student and the institution of higher education.

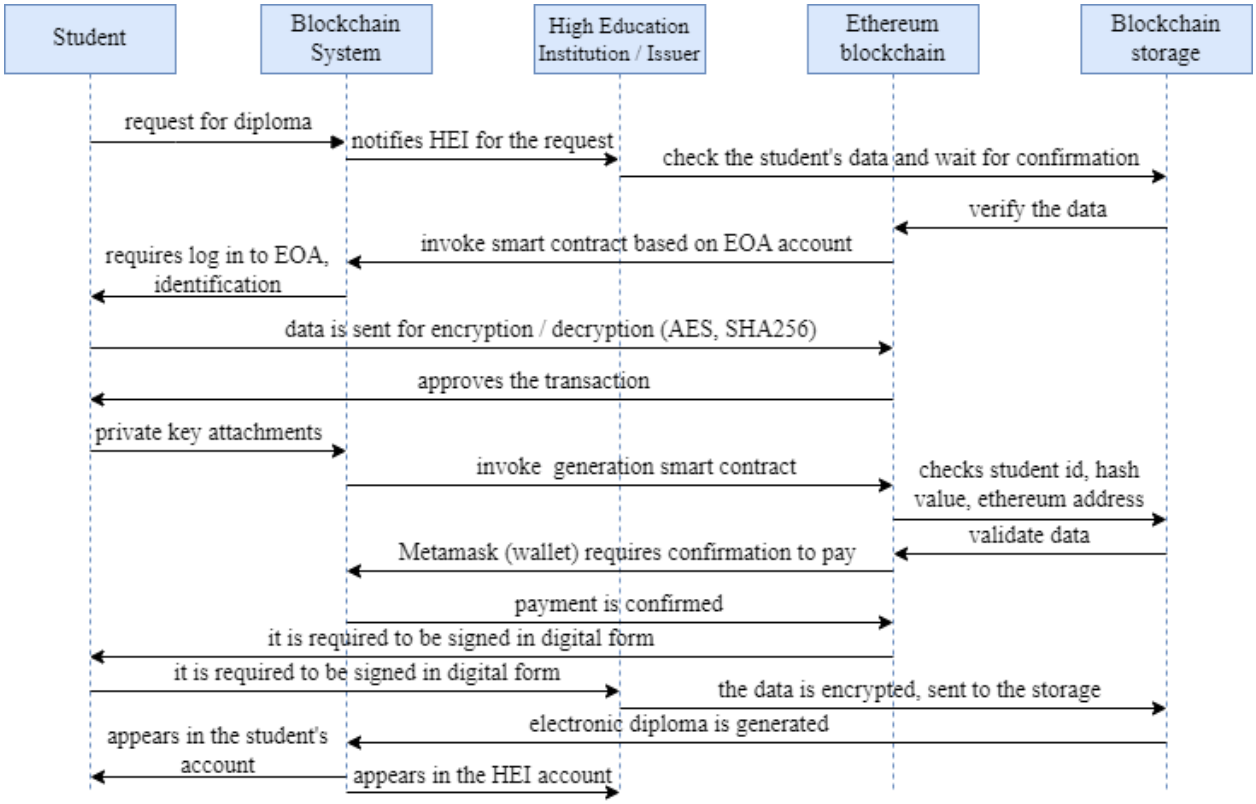

Fig 5. Sequence diagram of diploma generation form blockchain system

Based on the systematic literature review [1], which we have done regarding the implementation of blockchain in higher education institutions, it turns out that there are only few architectures, models and adequate literature that show the details of the implementation of systems or blockchain solutions in this direction. Although there is a considerable number of blockchain solutions, again there is no research that shows the success or failure of the ones or their practical implementation in any institution of higher education. It is characteristic that most blockchain solutions have used the Ethereum platform, Solidity for smart contract programming, and Truffle suite as a package with the tools for the development of the blockchain network infrastructure. Because we are in the phase of designing a system based on blockchain technology for the generation and verification of diplomas, after numerous analyses of the numerous challenges and limitations in terms of the practical implementation of the smart contract, it is very important to include all the details of implementation, to be as accurate as possible in programming the smart contract and overcoming the limitations so far. Only after we have implemented the initial prototype based on the architecture and the conceptual model we are developing, we will be able to have analyses and comparisons, respectively quantitative analyses of the performance of the system in real time, to test the validity of the same based on the defined architecture. The optimization and testing of the smart contract is of great importance for the creation of the blockchain

system for the generation and verification of diplomas. There are different smart contract testing techniques, such as Mutation testing, OYENTE, Zeus, Vandal, ContractFuzzer, GasSaver, SmartBug, while optimization has to do with the detailed analysis of the code during programming and the overcoming of useless and unnecessary elements with the aim of reducing the cost and execution time of the smart contract [21].

5. CONCLUSION

Regarding the application of blockchain in higher education institutions, we have done a lot of research, review, definition of architecture, use case diagrams and scenarios for the main actors in this system. We also designed the conceptual model, describing in detail the process of inserting and verifying diplomas, in the form of systematic architectures. This means that detailed analyses have been made in terms of the practical application of the system. Through this paper, after describing the importance of blockchain for higher education institutions, we described the importance of smart contracts and the challenges in terms of their application in higher education institutions. Through the paper we tried to make a comparison between centralized and decentralized systems, and then also the description of some decentralized databases that are towards development recently. And in the end, we paid special attention to the design through the class diagram of the blockchain system and in particular to describe the importance of the blockchain-based database for the blockchain system. The primary importance of this work is the creation of the architecture of how the smart contract execution process flows, the design and importance of the blockchain database in the creation of the blockchain system. The next challenge is the practical implementation of the system, quantitative analysis of the code's performance, including real-time analysis and results of the generation and verification of diplomas, which we will present in the next research.

REFERENCES

[1] Rustemi A., Dalipi F., Atanasovski V., Risteski A. (2023): A systematic literature review on blockchain-based systems for academic certificate verification, *IEEE Access*, Vol. **11**, pp. 64679–64696. DOI: 10.1109/ACCESS.2023.3289598

[2] Peng Y., Yang X., Zhou H. (2021): Blockchain technology and higher education: characteristics, dilemma and development path. In: *Proceedings of the 2020 4th International Conference on Education and E-Learning (ICEEL '20)*. Association for Computting Machinery, New York, NY, USA, 173–176. https://doi.org/10.1145/3439147.3439185

[3] Saleh O., Ghazali O., Rana, M. E. (2020): Blockchain based framework for educational certificates verification. *Journal of Critical Reviews*. **7**. 79–84. 10.31838/jcr.07.03.13.

[4] Hsu, C. S., Tu, S. F., Chiu, P. C. (2022): Design of an e-diploma system based on consortium blockchain and facial recognition. *Educ Inf Technol* **27**, 5495–5519. https://doi.org/10.1007/s10639-021-10840-5

[5] Gugnani P., Godfrey W. W., Sadhya D. (2022): Ethereum based smart contract for event management system, In: *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, Gwalior, India, pp. 1–5. DOi: 10.1109/CICT56698.2022.9997939

[6] Nguyen D.-H., Nguyen-Duc D.-N., Huynh-Tuong N., H.-A. (2018): CVSS: A blockchainized certificate verifying support system. In: *Proceedings of the 9th International Symposium on Information and Communication Technology (SoICT '18)*. Association for Computing Machinery, New York, NY.

[7] Geroni, D. (September 16, 2021): Top 12 Smart coctract use cases, [Online]. Available: https://101blockchains.com/smart-contract-use-cases/

[8] Awaji, B., Solaiman, E., Albshri, A. (2020): Blockchain-based applications in higher education: a systematic mapping study. In: *Proceedings of the 5th International Conference on Information and Education Innovations (ICIEI '20)*. ACM, New York, NY, USA, 96–104. https://doi.org/10.1145/3411681.3411688

[9] Rustemi, A., Atanasovski, V., Risteski, A. (2023): Identification during verification of diplomas. In: *The Blockchain System. 2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP)*, Ohrid, North Macedonia, pp. 1–5. DOI: 10.1109/IWSSIP58668.2023.10180241

[10] Cernian, A., Vlasceanu, E., Tiganoaia, B., Iftemi, A. (2021): Deploying blockchain technology for storing digital diplomas. In: *23rd International Conference on Control Systems and Computer Science (CSCS)*, Bucharest, Romania, pp. 322–327. DOI: 10.1109/CSCS52396.2021.00059

[11] Pandey, S. et al. (2022): Do-it-yourself recommender system: Reusing and Recycling with Blockchain and Deep Learning., In: *IEEE Access*, Vol. **10**, pp. 90056–90067. DOI: 10.1109/ACCESS.2022.3199661

[12] Zhu P., Hu J., Zhang Y., Li X. (2020): A blockchain based solution for medication anti-counterfeiting and traceability, *IEEE Access*, Vol. **8**, pp. 184256–184272. DOI: 10.1109/ACCESS.2020.3029196

[13] Ge, X. et al. (2023): Blockchain and green certificates based market structure and transaction mechanism of direct power-purchase for industrial users. In: *IEEE Transactions on Industry Applications*, Vol. **59**, no. 3, pp. 2892–2903. DOI: 10.1109/TIA.2023.3246966

[14] Kaur, J., Rani, R., Kalra, N. (2022): An automated liver disease detection system using machine learning and smart contract, *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*, Bhopal, India, pp. 1–5. DOI: 10.1109/CCET56606.2022

[15] Wang, Y., Su, Z., Xu, Q., Li, R., Luan, T. H., Wang, P. (2022): A secure and intelligent data sharing scheme for uav-assisted disaster rescue, *IEEE/ACM Transactions on Networking*. DOI: 10.1109/TNET.2022.3226458

[16] Rustemi, A., Atanasovski, V., Risteski, A. (2022): Overview of blockchain data storage and privacy protection, *2022 International Balkan Conference on Communications and Networking (Balkan Com)*, Sarajevo, Bosnia and Herzegovina, pp. 90–94. DOI: 10.1109/BalkanCom55633.2022.9900867

[17] Dneprovskaya, N. V., Bayaskalanova, T. A., Ruposov, V. L., ShevtsovaI, V. (2018): Study of digitization of Russian higher education as basis for smart education, *2018 IEEE International Conference*

"Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), St. Petersburg, Russia, pp. 607–611,
DOI: 10.1109/ITMQIS.2018.8524945

[18] Dwivedi, A. (2018): Digitizing academic delivery in higher education issues and challenges, *2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)*, Noida, India, pp. 179–182.
DOI: 10.1109/ETTLIS.2018.8485251

[19] Ehsan, I., Khalid, M. I., Ricci, L., Iqbal, J., Alabrah, A., Ullah, S. S., Alfakih, T. M. (2022): A conceptual model for blockchain-based agriculture food supply chain system, *Scientific Programming*, Vol. **2022**, 15 pages, Article ID 7358354.
https://doi.org/10.1155/2022/7358354

[20] Shoufan, A., Mohamed, F. (2022): YouTube and education: A scoping review, *IEEE Access*, Vol. **10**, pp. 125576–125599.
DOI: 10.1109/ACCESS.2022.3225419

[21] Rustemi, A., Atanasovski, V., Risteski, A. (2023): Design of the blockchain system for the generation and verification of diplomas, *2023 XXXII International Scientific Conference Electronics (ET)*, Sozopol, Bulgaria, 2023, pp. 1–6.
DOI: 10.1109/ET59121.2023.10279743