A Blockchain-based Architecture for Query and Registration of Student Degree Certificates

Antonio Welligton S. Abreu
Federal University of Ceará
Quixadá, Ceará, Brazil
siwelligton@gmail.com

Emanuel F. Coutinho
Federal University of Ceará
Quixadá, Ceará, Brazil
emanuel.coutinho@ufc.br

Carla I. M. Bezerra
Federal University of Ceará
Quixadá, Ceará, Brazil
carlailane@ufc.br

ABSTRACT

Blockchain is an emerging technology that has piqued the interest of researchers and industry today. Following the success of this technology in the financial market with virtual currencies, blockchain begins to be used by different domains such as government and education as it provides a reliable, scalable and unchanging distributed environment for conducting and storing transactions on a network. Higher education is a system with many challenges that could be solved by adopting this technology. Protecting data transactions that involve student degrees is one of the challenges considered by educational institutions. Thus, this paper presents a reference architecture proposal based on blockchain technology and a proof of concept to register and consult student certificates issued by higher education institutions. With this architecture, we intend to evaluate a new business model in this technology applied to the educational domain. As a proof of concept, we presented a prototype implementation using smart contracts based on the Ethereum platform and an evaluation made by experts in high education.

CCS CONCEPTS

• **Applied computing** → *IT architectures*; *Enterprise information systems*; • **Software and its engineering** → *Software organization and properties*; • **Computer systems organization**;

KEYWORDS

blockchain, architectures, smart contracts, education

ACM Reference Format:
Antonio Welligton S. Abreu, Emanuel F. Coutinho, and Carla I. M. Bezerra. 2020. A Blockchain-based Architecture for Query and Registration of Student Degree Certificates. In *14th Brazilian Symposium on Software Components, Architectures, and Reuse (SBCARS '20), October 19–20, 2020, Natal, Brazil.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3425269.3425285

1 INTRODUCTION

Blockchain technology is slowly integrating into different domains such as logistics, energy and health. One of the appropriate domains for blockchain adoption is higher education, where the principles of document authentication, transparency, immutability and trust are the key advantages that make this an attractive solution [10].

Most Higher Education Institutions (HEIs) have their own specialized system to maintain complete course and student records, where data is structured to be accessed only by the staff of an institution or students on restricted websites or dedicated systems to this end, having little or no interoperability. In general, the databases of these systems are hosted in a datacenter within HEIs, with restricted access to their IT professionals [21]. In HEI, the issuance and availability of the certificates is playing a critical role as it is strong evidence that the student has successfully completed the course [1]. Thus, it is important to prevent degree transcription fraud as well as a place to check the data issued accuracy by the institution. This can be done through blockchain technology as it allows high trust, disintermediation and verified transactions [19].

In view of this scenario, we can observe some challenges in access control and data security of students' diplomas: (i) there are several different databases with their own access restrictions, and security mechanisms can become a complicating factor in the information management issued by HEIs in regional or national territory; (ii) the need to not have third parties involved in the validating process of diploma data is also a challenge, as students need the HEI that issued the certificate to prove their training in the face of a selection of job vacancies or studies continuation; (iii) the fact that it does not depend on a manual administrative process or the availability of a HEI system to prove a diploma can be an improvement situation to be considered; (iv) each HEI is responsible for the safe storage of the students' diplomas data, and this can be a worrying factor, since the mechanisms used by each HEI may not be the most adequate to safely store the data for a long period of time, causing the loss of information; and (v) preventing forged documents is one of the great challenges of higher education, and it is necessary to have some mechanism of data verifiability to avoid falsification of certificates.

All of these problems mentioned can be addressed with traditional technologies, and they can be centralized or decentralized systems. In decentralized systems, we could have different architectural patterns and communication protocols of the conventional mechanisms for distributed systems (e.g. CORBA, .net, RMI, SOAP, REST). In distributed systems we have several components distributed by different servers, which can be heterogeneous, with different operating systems, manufacturers and processing capacities. In centralized systems this does not exist, as there is a single component which is the centralized system itself [11]. Even with a wide variety of mechanisms in these conventional technologies to deal with data security, there are still some challenges like the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SBCARS '20, October 19–20, 2020, Natal, Brazil
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-8754-5/20/09…$15.00
https://doi.org/10.1145/3425269.3425285

integrity and reliability of system transactions, where these technologies try to improve every day to avoid data security problems involved in transactions [26].

Blockchain is an alternative to address the mentioned problem scenarios, where transactions between trusted parties (student HEI, student, business, government and other HEIs) are required for the validation, storage, availability and security of data from diplomas. This technology can more adequately address the problems mentioned, as the solutions created through blockchain technology allow high trust, disintermediation and verified transactions [19].

In Brazil, the diploma registration process involves documents from students and HEIs that are essential to ensure the safety and validity of the legal acts to be produced by this process, and the diploma registration must be done in a specific book in the physical environment or electronic, at the discretion of each institution. Following the administrative act of ordinance number 1,095 by the Ministry of Education of Brazil (MEC), after registration it is necessary to publish an extract of the information on registration in the Official Gazette (DOU), paying a certain financial amount for the space used in the publication, where the responsibility for publishing the information in the DOU rests with each issuing HEI. The purpose of this new law is to give more credibility and security to the information of students' degrees [15].

In this context, this work proposes an architecture model using blockchain. This approach is about having a physically distributed infrastructure, but logically centralized, providing a single view of the truth about the students' diploma data for all interested parties. The main objective of this work is to provide an environment with credibility and security through blockchain technology to publish diplomas information of higher education students, as well as the consultation for diploma data validation. With this technology it is possible to reduce the risks of loss of information, in addition to saving paper, reducing management costs and preventing forged documents, because through the data verifiability principle, it is possible to avoid certificates falsification. The problem addressed in this work is to build an approach that uses blockchain technology and to identify what are the main technological challenges that educational institutions may face to innovate in the area of management of diploma data using blockchain technology.

The main contributions of this work are: (i) a reference architecture for blockchain; (i) a prototype of an application in the educational domain that uses blockchain resources; and (iii) an evaluation from technical and user view about the solution.

The article is divided in the following sections: Section 2 presents the Background; Section 3 presents the related work; Section 4 describes the methodology; Section 5 shows the proposed architecture; Section 6 describes aspects of the prototype development; Section 7 shows results and discussion; Section 8 presents the threats to validity; and finally, Section 9 presents conclusions and future work.

2 BACKGROUND

2.1 Degree Registration and Access in Higher Education

According to the Brazilian Ministry of Education (MEC), more than 1 million higher education students graduated throughout Brazil each year, and in 2018 about 1.2 million graduate students [16]. Certificates that students receive upon completion of a course prove their education, performance and they are an important reference for entry into other institutions or jobs [6].

With the technology advancement in recent years, the need for student and institution data protection has grown. Graduates prove accuracy through certificates for use in selection processes. However, they often end up losing the printed certificates, and a copy must be requested from the organization issuing this document, which may take a long time because of the administrative processes for validating the data for issuing the certificate. On the other hand, ordering an electronic copy can save paper and time. However, because of this convenience, the falsifying certificates likelihood is high, so educational institutions and companies can not quickly validate the received documents [6].

There are vital points regarding systems responsible for controlling educational data, such as data standardization, storage location, security, and how to filter, analyze, protect, and share this data. Because of these problems, higher education institutions keep track of complete student data indefinitely for legal reasons, depending on a country's policy [21].

In Brazil, the administrative act of Ordinance No. 1,095 was recently issued by MEC, stating that public and private HEIs that have the prerogative to register the diplomas issued by them, should publish an extract information about the registration in the Federal Official Gazette (DOU: means of communication by which the National Press has to make public any and all matters about the federal scope). They shall also maintain a database of diplomas registration information to be made available on the HEIs website after proper registration in DOU [15]. With these new rules, the Brazilian government intends to improve the internal processes of the institutions, giving information agility and security to deliver a better student experience.

2.2 Blockchain

Blockchain is a block sequence that contains the complete transaction log as a public book, maintained by multiple nodes in a network [3]. Each node contains the identical copy of this ledger, each block being a logical sequence of transactions, which are permanent, transparent and unchanging records [20]. Each block contains a timestamp, the hash value of the previous block (parent), and a nonce, which is a random number to verify the hash. This concept guarantees the integrity of the entire blockchain to the first block (genesis block). Hash values are unique and fraud can be effectively prevented, since changes to a block in the chain would immediately change its hash value [2]. Figure 1 represents a blockchain where the newly validated block points to the immediately preceding generated block. Each block in the chain confirms the integrity of the previous one, and all the way back to the first block [3].

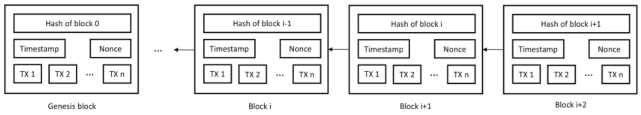


Figure 1: Example of blockchain and its blocks [2].

Blockchain technology can record transactions between two parties efficientl, verifiably and permanently, eliminating the need for third parties. In addition, the availability of all completed transactions for all nodes makes a blockchain-based system more transparent than centralized solutions. Transactions are validated by member nodes using a consensus protocol, which ensures that all nodes have an identical copy of the blockchain. A new block is considered verified only after most member nodes vote as true and reliable using the consensus protocol [4].

Implementing blockchain requires a lot of resources. An important parameter for analyzing blockchain is the block itself, by its size, time required for mining, and cost. Mining is the process of validating a block in blockchain. Something important is that it takes a lot of computational power to become a blockchain miner. Miners are usually rewarded (e.g. Bitcoin). Mining can affect all system performance and it is a critical concept [18].

The question of determining what type of blockchain and what configuration to use in developing solutions has presented a major obstacle in decision making between manufacturers and system architects. Even though they have a structure to address these issues and comprehensively explain the technical design and business application considerations, they fail to address common decision making to see if a blockchain solution is viable, and if so, what kind blockchain should be implemented. Each implementation of a blockchain requires careful analysis to decide based on the individual characteristics of the application [17].

3 RELATED WORK

This section presents works involving applications in the field of certificate management in higher education and blockchain. Recently, some universities and institutes have applied blockchain in education, most of which have used this technology to support academic management.

Cheng et al. [6] proposed an Ethereum decentralized application for digital certificate issuance. Its purpose was to save paper, reduce management costs and prevent counterfeit documents by providing accurate and reliable digital certificate information. The digital certificate issuance procedure in this system involves the generation of a paper certificate electronic file that accompanies other related data, after which a hash value for the electronic file is calculated. It stores the hash value in the system blockchain, where it will create a QR code and a query related to the code in the blockchain. The QR code is fixed on the paper certificate to verify authenticity by scanning a mobile phone or by consulting the application's website. The system not only increases the credibility of certificates, but also electronically reduces the risk of information loss. The differences in the work proposed by [6] are related to the way of consulting and controlling the data insertion in the blockchain, as this work does not consider the government entity in the control of HEIs that can insert data in the blockchain. Thus, an unregulated HEI could interact with the blockchain without any hindrance. This situation will be addressed in the proposed work.

Taufiq et al. [19] described a proposal for storing postgraduate degrees using Hyperledger Fabric and blockchain network protocols to provide security and transparency of undergraduate student data for the purpose of preventing certificate fraud. Comparing with the proposed work, this work did not create an application to validate the system model and also the number of nodes to establish the consensus of transactions in a system is small, as it is a permissioned blockchain network. The proposed work will use Ethereum's public network (without permission), where it has a large number of nodes for the consensus of transactions.

Kanan et al. [12] developed an electronic authentication system for documents using Ethereum, aiming integrating official information of all types of documents can be issued by a university. The created system was for a specific HEI with restrictions to insert data into the blockchain. The structure of the created smart contract is limited to the HEI scenario, different from the proposed work that suggests a universal structure for all HEIs in a region.

A good overview of current trends and existing implementations of certificate control blockchain can be found in [5] and [9]. The first work presented some ongoing projects as a form of market research to see the usefulness of blockchain, different from the proposed work that will present an architectural model as an approach. The second work used the Bitcoin blockchain, requiring the installation of an application to use the created system, unlike the proposed work that will use the Ethereum blockchain and it will not be necessary to install anything to use it.

Table 1 presents a comparison between the related work and this proposal.

Table 1: Comparison between related work: Work - work reference; Application - if an application has been created; Public - if used a public blockchain; Query - if there was a public query; Registration - if a registration was required

Work	Application	Public	Query	Registration
Cheng et al. 2018 [6]	Yes	Yes	Yes	Yes
Taufiq et al. 2019 [19]	No	No	Yes	Yes
Kanan et al. 2019 [12]	Yes	Yes	Yes	Yes
Chen et al. 2018 [5]	No	No	No	No
Durant et al. 2017 [9]	Yes	Yes	Yes	Yes
Our Proposal	Yes	Yes	Yes	No

While new use cases and business models may emerge with blockchain technologies, several challenges have become seeming when implementing a prototype. Most related work provide a public consultation to validate student certificate data, but a prior registration in the system is required to perform the consultation, unlike the proposed work, which any entity can consult the data without the need for a registration in the application.

According to ordinance 1.095 of MEC, all HEIs in the country will need to make information available via the Web to consult the diplomas. Thus, for each HEI, to create an infrastructure that has security, storage, availability and access control, it may be complex, especially for small HEIs. So, using a blockchain-based solution we can have these requirements met with less effort than a traditional solution.

4 METHODOLOGY

At this moment, an intention of this work is to carry out a proof of concept on the idea of validating the use of blockchain in the educational field.

The methodology of this work consists of the six steps, described as follow. (i) *Proposition of a reference architecture for blockchain:* Activities for proposing a reference architecture for blockchain in layers, with the intention of directing and standardizing development; (ii) *Development of a prototype for an application in the educational domain:* Specification and development of a prototype for the educational domain with blockchain to assess the applicability of the technology; (iii) *Design of a usage scenario:* Activities for the design of a scenario in the educational domain, for the consequent application of the prototype; (iv) *Technical validation of the prototype:* Mapping of prototype characteristics with the criteria defined by Ciccio et al. [8] for monitoring the business process with blockchain; (v) *Validation of the prototype for the applicability of the blockchain:* Validation of the applicability of the blockchain with the educational domain from the viewpoint of viability through the steps described in Pedersen et al. [17]; and (vi) *Prototype validation with user:* Prototype validation activities with a user through interviews, with pre-established questions, aiming to compare the use of the tool with the current usage process.

5 REFERENCE ARCHITECTURE AND PROPOSED SOLUTION ARCHITECTURE

Viewing blockchain as a software component helps to understand important architectural impacts on the performance and quality of attributes such as security, privacy, and scalability. As a component, blockchain has unique properties and limitations, being complex, having network-based software components that can provide data storage, computing and communication services [26].

The reference architecture proposed in this work was based on [14][22][25][26]. Its elements are described as follows: **Application:** This layer includes applications that use resources from a blockchain. Such applications may include business logic and data, in addition to having a graphical user interface; **API:** The blockchain can interact with elements from the outside world. This communication is carried out through APIs, both for consultation and for writing data; **Blockchain:** The blockchain layer is responsible for storing and sharing data, in addition to executing smart contracts. It is subdivided into the following components: (i) Smart contracts: A smart contract is a program created by the user, deployed and executed on the blockchain, and can represent business rules. They can also be implemented as part of a transaction; (ii) Transactions: This component has implementations for the generation and validation of the blocks and their order; (iii) Keys: Component for key management enables blockchain participants to own their private keys, used for digitally signing transactions; (iv) Ledger: Component that represents the blockchain distributed ledger; **Conventional database:** Databases to assist the storage of data that will not be stored on the blockchain, either by large size, either due to scalability or privacy; **Consensus algorithms:** This layer allows implementations to generate the order of the blocks and validations by the network nodes, with defined algorithms; and **Network:** Layer responsible for communication between nodes, including discovery, propagation of transactions and blocks.

Figure 2 shows the proposed blockchain reference architecture, as well as its layers and components, and it is the basis for creating the proposed architecture.

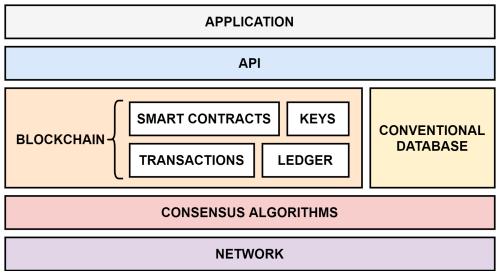


Figure 2: Proposed blockchain reference architecture

The proposed blockchain architecture is different from the usual architectures in some ways. The manner the solution is built and maintained is different from traditional ways. The application executes the transactions using a hash or proof of work algorithm to compute the information through a consensus process in a network with several nodes, ensuring the integrity of the information stored in the blockchain. This is a complex process, but currently there are several tools that encapsulate all this complexity, facilitating the creation and maintenance of this type of application. With that, we can have a system with security, storage, high availability and access control with little development and maintenance effort. In building the blockchain-based approach, we realized that there are several specific components of the blockchain area. With this, it was necessary to propose and use a new architectural model as a reference.

As defined in [26], with this reference architecture, information about the world outside the blockchain can be provided by specific applications, where it usually adds this information to the blockchain through transactions. Key and token management allow control of permissions and access to a blockchain. Part of an application can be implemented within the blockchain component using ledger and smart contracts. Blockchains can be used as software components, which can provide data storage, computing services, communication services, and management functions. For blockchain-based systems, the main architectural decisions are about which parts of the data to include or not in the chain. However, the amount of computational power, data storage space, and access control on a blockchain can be limited. Thus, parts of an application implemented outside the blockchain component can host offline data and application logic [26].

The architecture proposed in this paper aims to provide a credible and secure environment through blockchain to publish certificate information extract from higher education students, as well as the consultation for validation of certificate data. With this technology, it is possible to reduce the information loss risks and through the data verifiability principle, to prevent certificates falsification.

Figure 3 presents the architecture divided into three layers: (i) Application layer, called Educ-Dapp responsible for interacting with entities outside the system; (ii) API layer, composed by the technologies required to create blockchain access; and (iii) Blockchain layer, where the educational data and smart contracts reside.

Observing in the reference architecture (Figure 2), we have a consensus algorithm layer, in which it offers a set of rules that establishes how the blocks should be built in the blockchain layer.

Participation in the consensus-building process can be supported by several different structures: Proof of Work, Proof of Participation, Tolerance to Byzantine Failures. Regardless of the method used, the network nodes actively participate in this consensus process. We also have in the reference architecture the network layer, where applications and business processes are determined from their decentralized structure. For the connection to the network to be established, a node must be able to process messages specific to that network and affect its status. These consensus and network algorithm layers are incorporated into the logic of the proposed architecture, being layers very focused on infrastructure and networks, so they will not be the scope of this work.

The first layer is composed by the decentralized application Educ-Dapp, responsible for direct communication with student entities, educational institutions, companies and government, the main stakeholders defined in this work. The educational institutions and government entities are the only ones that can enter data in the blockchain. The creation of the smart contract will automatically generate a login for the government to enter the system, where it will be responsible for entering the basic data that will be the name and national registration of the legal entity (CNPJ) of the regular educational institutions that can issue certificates. Thus, these data will be used for validation at the moment an educational institution is making its registration in the application, that is, it will only be able to register if the government has previously informed its basic data in application. Otherwise, it indicates that it has no authority to issue certificates in the country. The student entities and companies can consult the certificates data without the need for registration, just have the student's individual registration number (CPF).

The second layer consists of the technologies responsible for integrating the application front end with the blockchain layer. To create a blockchain application, it is required to use specific settings on the front end in order to communicate with blockchain. There are specific tools for using this technology. This layer corresponds to the API layer of the reference architecture, providing technologies to allow developers to access the necessary resources and integrate them with applications created to interact with the blockchain.

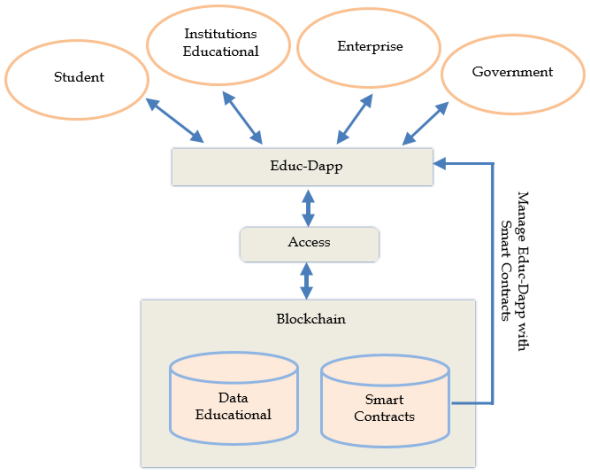


Figure 3: Proposed architecture

The third layer of the environment is formed by the blockchain that stores the students' educational data, which will be used for validation at the time of consultation by external entities. This layer also stores the smart contract, which defines the rules for how educational data is stored and accessed. These smart contract rules are informed within the Educ-Dapp application to manage query requests and insert educational data into the blockchain.

This approach works as a new business model for HEIs, students and the government, having its own characteristics and ways of dealing with the storage and access control of student certificate data. This business model uses blockchain technology to create an environment where HEIs can store student certificate information in a secure and distributed manner without the need to setup an internal infrastructure. Any entity external to the system can consult the database of diploma records stored on the blockchain. This approach will take advantage of the unalterable property of the blockchain and the data verifiability principle to avoid certificate forgery and to optimize the query time for the veracity of the data.

6 PROOF OF CONCEPT

The Ethereum platform was selected to create a proof of concept of the proposed architecture, being a more generic blockchain with smart contracts, and expanding transactions to computational operations that can be programmed to execute certain rules. Smart contracts function as scripts stored in the blockchain, having a unique address. An intelligent contract can be triggered by addressing a transaction to it, where it then executes independently of the way it was written, on any network node, according to the data that was included in the transaction trigger [7].

There are several networks on the market today that can work with smart contract creation. Ethereum has been selected for prototype creation as it is currently well suited for smart contract execution in the context of blockchain [13]. It has the Ethereum Virtual Machine (EVM) [24], where the proof of concept application (smart contracts) was created, which works exactly as programmed, without any possibility of censorship or fraud, as the contract is unchanging. The created prototype uses a public blockchain without permission from Ethereum, that allows any external entity to consult the data stored in the blockchain.

Some Ethereum platform-related technologies used for prototype development were: **IDE Remix:** Online IDE for smart contract development; **Solidity:** Ethereum's own language for smart contract development; **Web3 API:** web3 documentation JavaScript Dapp API used for decentralized application development Educ-Dapp; **Metamask:** Browser plugin (Chrome or Firefox) to access the Ethereum platform, where it works as an ether wallet and also browses the Ethereum network without having to have a copy of the blockchain installed in the local environment; **Ropsten:** Ethereum's public testing network to validate the prototype; **Infura:** website to interact with smart contracts because we need to be connected with a node, which is the gateway to the Ethereum network; and **Etherscan:** a tool for exploiting a blockchain, allowing you to analyze transactions on the Ethereum platform as a way to aid validation during application development.

Solidity is a high-level, strongly typed language with libraries, inheritance, and object orientation, having similarities to JavaScript.

Upon completion of smart contract programming, a compiler named solc was required to transform Solidity code into the contract bytecode, which is then interpreted by EVM. Then the compiled statements are deployed to an Ethereum blockchain.

For this work, we developed a smart contract with two control structures, one for storing educational institutions and other for storing student diploma data. There is a method that records the degrees in the blockchain, receiving the defined data to be stored, and there is a method for retrieving the degree data for validation and query result display in the Educ-Dapp application.

Metamask plugin was used for smart contract implementation. It acts as an intermediary to perform transactions on the blockchain network. In this prototype, we used the Ropsten public network, classified as testnets, which use fictitious ethers to perform operations, and could try different features before publishing contracts on the mainnet. Through an account created in this plugin, the user authorizes or not the transactions on the blockchain, being able to manage all transactions through this account. After implementation of the smart contract, the front end application (Educ-Dapp) integration with the blockchain layer was developed. For front end programming, we used the Visual Studio Code IDE and the web3 JavaScript Dapp API documentation.

Figure 4 displays an overview of the execution flow of the technologies used during transaction executions on the created prototype. In the application, entities can use the blockchain to send and retrieve data. The goal of the front end is to provide an interface for the user to interact with the created smart contract. The front end is also informed of the contract code in JSON format, where from the transaction confirmation through the Metamask plugin, the Web3 API recognizes the Solidity code, that is in JSON format and initiates communication with Ethereum's testnet Ropsten Network through of the node created on the Infura website. To query the data in the blockchain is not necessary to use the plugin Metamask, because after the data entered in the blockchain any entity can consult without the need to perform some form of registration or installation as a prerequisite for access to data validation.

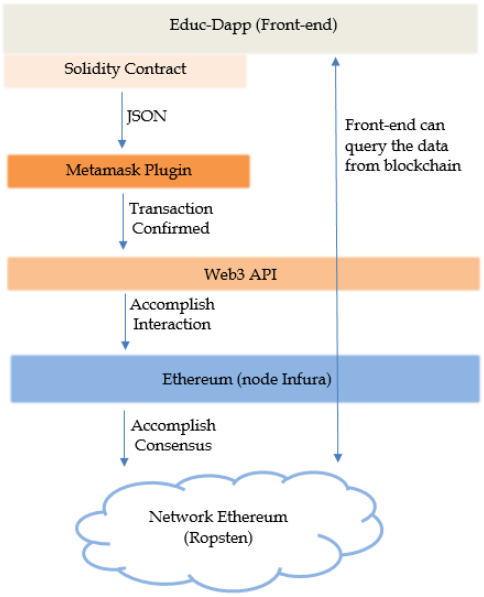


Figure 4: Execution flow of the used technologies.

7 RESULTS AND DISCUSSION

7.1 User Validation Scenario

As it is not necessary for application users to understand blockchain technology at a low level (programming and infrastructure), this validation is about the business logic of the registration of diplomas, in which all interaction is performed with an interface similar to a traditional system.

To illustrate the architecture, a scenario was defined for a private higher education institution to perform the validation, where the participating key users were from the institution's degree issuing sector. The information needed to access the application was passed on to the institution, as well as the basic information to use its features. The blockchain technology was also explained in general to let them know that the application uses a nontraditional solution.

The educational data entered by HEI in the blockchain is based on the administrative act of Ordinance No. 1,095 by MEC. It considers the following data for the degree registration in the application: (i) Name of the institution; (ii) CNPJ number of the institution; (iii) Name of graduate student; (iv) CPF of the graduate student; (v) Name and MEC code of the college; (vi) Date of entry into the course; and (vii) Course completion date. All the data used in the validation scenario are fictitious, i.e. a scenario with data similar to the real environment to validate the created prototype.

Validation was performed by two participants of Centro Universitário Católica de Quixadá, Ceará, Brazil, each of whom has 10 years of experience in higher education. They are responsible for the entire process of issuing the private institution degrees. It was made clear to them that the government is registering the educational institutions that can use the application in the proposed architecture. The login that makes the institutions registration is generated at the contract time of creation and passed on to the government.

Institutions make an application registration to be entitled to register the diplomas, where the application makes a validation verifying if it has the registration informed by the government. If so, the system confirms the registration made by the institution. This prevents anyone from registering with the application. The other entities only have the permission to consult the data of the diplomas, searching by the student's CPF.

In this validation, there was not a government participant, so, initially the institution was registered using the government profile, to be authorized to perform the self-registration and have permission to register diplomas in the application. Then, the participants register the institution with the basic data, name and CNPJ. After successful registration, users logged in to start the degree registration process.

To publish (perform the registration in the application) the extract of the diploma information of some students, to perform the validation consultation of the diploma data, 10 cases (number considered sufficient by the participants) of tests were defined by the participants to understand the dynamism of the application. Figure 5(a) shows the moment when one of the cases was executed.

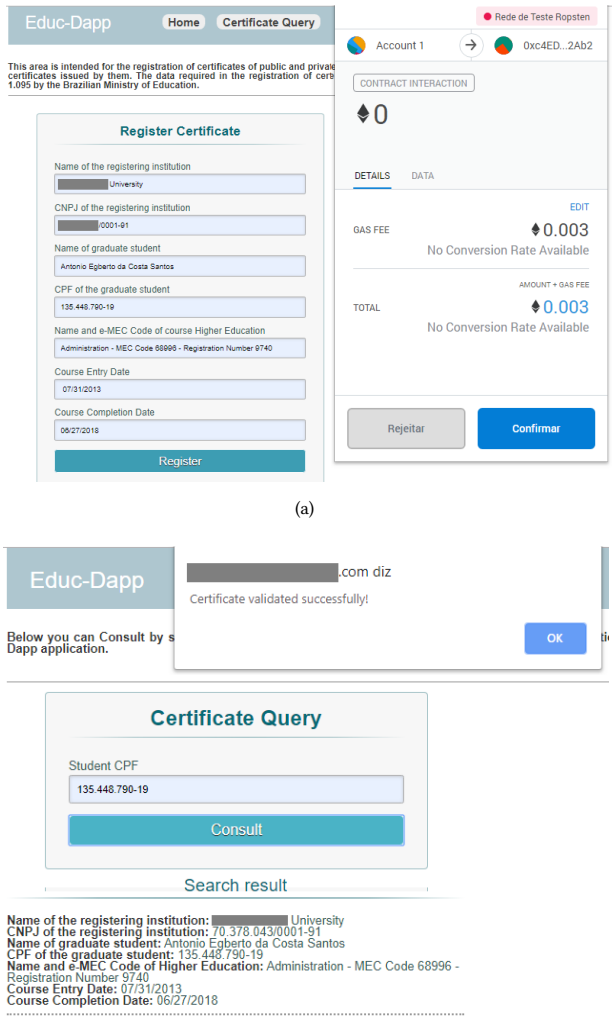


Figure 5: Application screens (a) Degree registration by IES e (b) Blockchain degree search result

Note that there was also a confirmation request from Metamask, as this is a data insertion on the blockchain.

After execution of the 10 test cases, the data became available for consultation. Figure 5(b) shows the moment that the query was performed using the CPF registered in Figure 5(a), where the application then searches the blockchain for the existence of a degree with this CPF. If so, the system displays a message confirming successful validation and displaying the data that was recorded.

7.2 Technical Validation

The use of blockchain in application development poses several technical challenges. Ciccio et al. [8] described some comparison criteria for monitoring the business process with blockchain, based on smart contracts (reflecting monitoring capabilities), oracles (identifying data sources) and data monitoring (balancing between data inside and outside the chain). There are challenges for each of these criteria.

As challenges on smart contracts we have: (i) Monitoring Transparency: anybody can see the logic and data behind monitoring, data required for smart contracts must be provided and confidential data should not be shared or be protected; (ii) Observability: to access monitoring data, smart contracts must implement mechanisms to expose them, for instance, private variables can not be easily accessed; and (iii) Lack of reactivity: smart contracts can not directly invoke external services.

For challenges on oracles: (i) Time management: the blockchain lacks the notion of time, and timers can not expire by themselves, an external triggering event is always required; (ii) Reliability: oracle breaks the decentralizsed trust and to compensate, several oracles should be used; and (iii) Flexibility: oracles are bound to smart contracts because a compromised oracle can not be replaced during execution, nor new capabilities can be introduced.

Finally, as challenges for data management: (i) Data quality: the quality of the data influences the monitoring results, for example, poor quality data sources compromise the monitoring and once stored, incorrect data cannot be changed; (ii) Data size: the cost of the blockchain is proportional to the amount of data stored, for example, in public blockchains, monetary cost (cryptocurrency), in private blockchains, overhead on the platform, and also storing data off-chain; and (iii) Side effects: most blockchains are prone to forks, and this may cause contradictory information or interoperability issues.

These technical challenges (which can become research challenges), can be categorized according to three main perspectives [8]: Design rationale (need for a complete logic of the project behind the realization of the monitoring infrastructure), Lack of capacity (need to address the lack of resources that the use of blockchain infrastructure brings) and Trust management (demand for a trust management policy with information and actors).

Table 2 shows the technical challenges with the perspectives analyzed in relation to the proposed work. The answer options were Yes, No and Partial to analyze the behavior of this work proposal and understand the benefits and obstacles in the development of blockchain-based applications. When answered with Yes, we indicate that the challenge is aligned with the need defined in the perspective to have a desired behavior in the created application. When answered Partial indicates that the challenge was not fully addressed according to the perspective of the perspective. Finally, answered with the option No indicates that the challenge was not dealt with according to the perspective.

For Monitoring Transparency, the perspectives were answered with Partial, Yes and Partial due to the respective reasons: sophisticated logic was not applied to deal with all points to verify possible deviations in the approval of smart contract transactions; it was necessary to address the lack of blockchain infrastructure resources to make changes to data after being stored on the blockchain; a complete trust management policy was not created between the actors of the application. Observability was answered Yes, No and No for the reasons that a complete logic for application was designed, it did not face difficulties of lack of blockchain resources and it was not necessary a policy of trust management between the actors. For Lack of Reactivity, Partial, Partial and No were answered, as

Table 2: Perspectives for analyzing the challenges: DR - Design Rationale, LC - Lack of Capability, TM - Trust Management

Aspect	Challenge	DR	LC	TM
Smart Contracts	Monitoring Transparency	Partial	Yes	Partial
	Observability	Yes	No	No
	Lack of Reactivity	Partial	Partial	No
Oracles	Time Management	Partial	Yes	No
	Reliability	No	Partial	No
	Flexibility	Yes	No	Partial
Data Management	Data Quality	Partial	Partial	Partial
	Data Size	Yes	Yes	Yes
	Side Effects	No	Yes	No

some exception handling mechanisms were implemented and we identified the lack of some resources to create monitoring responses for the application users.

For Time Management, the answers were Partial, Yes and No due to the lack of resources in the blockchain infrastructure to measure the exact time of a transaction, even so it was possible to verify the approximate time of transactions in the blockchain environment without the need for a policy trust between users. For Reliability, No, Partial and No were answered, as no logic or trust policy was created between users to validate the data before being stored on the blockchain, yet the infrastructure offers the possibility to create exceptions before the data is stored on the blockchain . Flexibility was answered with Yes, No and Partial, as the application actors interact directly with the system based on a management of the actors' information and no blockchain infrastructure limitations were identified at this point.

For Data Quality, it was answered with Partial, Partial and Partial due to the process of inserting data in the blockchain depending on the trust in the actors, even so some exception handling mechanisms were created to reduce the risk of inserting wrong data in the blockchain. For Data Size, it was answered with all Yes, because the smart contract was created optimizing the information to decrease transaction costs, as well as the standardization of the size of the data inserted in the blockchain was defined as a management policy. For Side Effects it was answered with No, Yes and No, as no logic was developed to avoid ramifications in the blockchain, that is, two or more blocks pointing to the same predecessor, in which this is considered a problem and no resources were identified in the infrastructure to avoid this situation.

7.3 Viability of the Blockchain Use

For a viability analysis for blockchain, an interview was conducted with the same participants of the user validation scenario, to analyze the proposed model, seeking to investigate justifications to support architectural decisions about whether to use a blockchain or some conventional system. Two approaches were used: the ten steps proposed in [17] and how to use the application compared to the current process.

7.3.1 Ten-step Path for Blockchain Validation. For this analysis, we used the ten-step path for deciding whether or not to use blockchain technology in the desired application [17]. The ten steps consist of questions that must be made before implementing blockchain to ascertain whether the use of the technology is feasible. Figure 6 presents the ten steps. Only questions 1, 2, 3, 4 and 7 were applied to educational specialists. Such questions are more related to business logic, and the others are technical and interview participants would have difficulty or would not know how to answer. As a general rule, [17] recommends a blockchain is viable if five questions of the applied questionnaire are answered as "Yes". Even so, professionals need to carefully balance various requirements for each application case individually.

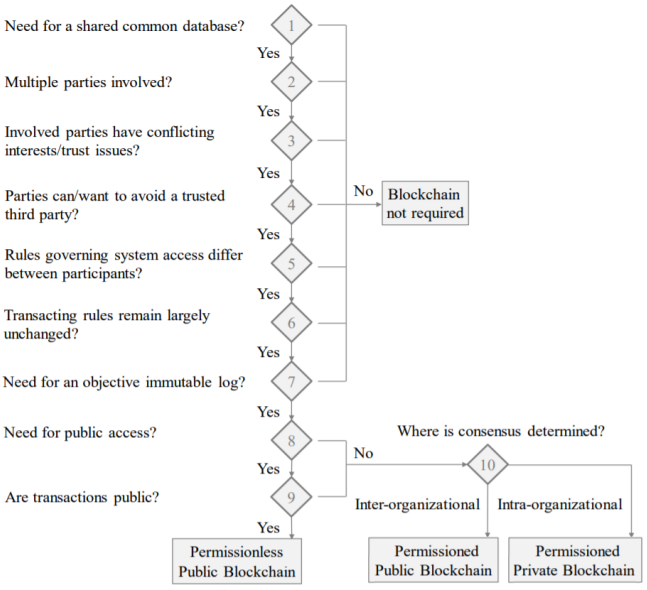


Figure 6: Summary of the blockchain decision path [17].

Need for a shared common database? *Justification: currently, the registration is made in the DOU, considered a central Brazilian government database.* Yes. Given that each institution or company have the need to validate degree data, there is nothing better than a single and shared basis;
Multiple parties involved? *Justification: The more individuals and institutions involved, the more complex data management is.* Yes. Academic Secretariat, Diploma Registration, Student, Companies and Public agencies.
Involved parties have conflictings interests/trust issues? *Justification: Because each institution and the government itself has their own databases keeping copies of the information, this can lead to information conflicts.* Yes. There is a need for validation of the authenticity of the data, both by the registrars as by those who consult this information (student, companies, public agencies etc.). By having several entities involved in the process, you may have conflicts of trust, as they are not known. With that, the blockchain can provide a formality for executing transactions and storing data.
Parties can/want to avoid a trusted third parties? *Justification: Most institutions today use third-party systems to validate information.* Yes. Once the information has been validated, there is no need to go through third parties. Third parties involved in the validation process (having a system from another company) can impair the control and access to the information in the diplomas. With

blockchain it would no longer be necessary to rely on third party systems.

Need for an objective immutable log? *Justification: Conventional databases are subject to possible security breaches known in the literature.* Yes. It is good to keep a log of the operations, either by consulting or publishing this data. Possible audits become simpler through unchanging logs. The data saved on the blockchain is immutable, this provides a mechanism for future audits.

7.3.2 Usage Assessment Compared to Current Process. To also evaluate the use of the Educ-Dapp application, compared to the current process required by MEC Ordinance No. 1,095, an interview questionnaire was elaborated, in which participants answered according to the experience of using the application and the information provided in begin of validation.

How difficult is it to use the Educ-Dapp system compared to the current process? *Justification: It is interesting to assess the level of difficulty between the Educ-Dapp system and the current process.* Compared to the current process, Educ-Dapp seems more simplified because to register certificates in DOU there is a specific format and configuration. Further, for registration in DOU is also necessary to configure the degrees in a third party system, as it is responsible for external consultation function.

Do you think the idea of Educ-Dapp helps improve the degree registration process? *Justification: It is interesting to evaluate the advantages of applying to the internal process of the institution.* With the Educ-Dapp application, the process would be in a single system, i.e., it is believed that would simplify every current process we perform.

Do you believe in the benefits pointed out by blockchain technology that underpins the proposed solution? *Justification: It is interesting to evaluate the level of confidence that this new technology gives to the validation participants.* As this is a new technology, we do not know whether the benefits actually occur in practice.

Is there any suggestion for improvement? *Justification: It is interesting to evaluate if participants see the limitations of blockchain technology.* This issue of creating a cryptocurrency wallet in the Metamask plugin to pay for transactions is difficult to align with the institution's internal payment processes, and it is not known how to handle this financial cost in accounting. This is a legal issue that should be considered before using this solution in the market.

Questionnaire responses point to advantages in using the proposed model, but institutions and policy makers must consider the challenges related to security, privacy, cost, scalability and availability before adopting blockchain technology.

7.4 General Discussions

The reference architecture implementation allows a standardization for applications development that integrates diverse services, of various technologies (web, mobile, microservices, etc.) with blockchain. Thus, the benefits of using blockchain can be shared for several services. One of the ideas of the reference architecture is to identify the most commonly used layers, organize them for a better understanding and facilitate their use, instantiating for applications.

Regarding quality criteria for applications that use blockchain, it is still necessary to explore further, especially in relation to interoperability. As systems from different platforms will be communicating, exchanging data and generating new information, the impact of non-functional requirements must be assessed with focus on the quality of the customer's product. This may justify the adaptation of traditional systems with solutions that use blockchain.

As an initial assessment with users working in the educational area, it was noticed there are benefits that can improve the quality of activities. There is also an impact on cost, as there is a need to further assess the financial impact of using blockchain in applications, if using a paid network.

In the context of higher education, smart contracts on blockchain networks will enable secure and autonomous certificate enrollment and availability solutions, creating transactions that will be executed only after specific requirements have been completed. This not only allows a greater automation, scalability, and cheaper transactions (no third party is required to oversee transactions), but these smart contracts can also prevent overwriting by people who want to use the data for their own benefit. Information is shared on a decentralized, encrypted-protected network, which means it is very difficult to compromise network security [24].

8 THREATS TO VALIDITY

This research can be affected by different factors which can invalidate its main findings [23].

Internal Validity. Internal validity is concerned with the validity within the given environment and the reliability of the results. The network environment can vary (latency, execution time of inclusions and transactions queries, block validation, disk space, etc.). This may not be good if the application needs immediate results. To minimize, several tests on the application can be performed to analyze quality attributes to suit the application and the environment. Another situation is that once the data is registered on the blockchain, even though the process is all correct, the data may be wrong. As the blockchain is immutable, this incorrect data is recorded. To mitigate this situation, a new correction record must be added, and the application must handle this case.

External Validity. External validity is the extent that the obtained results of a case study can be generalized to other relevant research scenarios. Only one institution in the experiment was used in the study. In addition, the involved professionals discussed the process they know about the institution in question. Other institutions have different business processes, with different applications, for the execution of certificate registration. Additionally, few people evaluated the application and few were interviewed. Another point is how scalable the prototype is. The approach can be easily used in any institution regardless of the number of students, as it follows the standard defined by MEC ordinance 1,095. However, validation was carried out in a single private HEI with around 3000 students enrolled and an average of 300 graduating students per semester.

Conclusion Validity. Conclusion validity is concerned with the relation between treatment and outcome. The applied questionnaire and interview had few, and subjective, questions. The results were much of the interviewees' view, which were few in quantity. To minimize this threat, better structured questionnaires

and interviews, with business and technical issues should be designed, with objective and subjective questions. And apply with more professionals, to obtain more significant results.

Construct Validity. Construct validity is concerned with the relationship between theory and observation. This work considered only the opinion of people from only one institution. Moreover, they did not have the technical knowledge of blockchain, making deeper responses impossible, as they only had knowledge in the educational institution's processes. The used data was fictitious and not production. Additionally, the application was not made available for use in production, only for this research. There is a need to evaluate with real data, with large quantities, impacts on quality attributes, with multiple users, to really prove the effectiveness of the application and whether it is worth investing in blockchain.

9 CONCLUSION

A wide variety of applications are rapidly emerging with blockchain technology, but in education it is in its early stages, where various industry organizations and academia are seeking the benefits of this technology for education.

According to the results presented in the validation scenario, the proposed approach based on blockchain technology could provide a viable alternative to store and control access to student certificate data, as it uses a secure platform to share data and increase trust and transparency. in consulting the veracity of the informed students diplomas. Regarding privacy, validation follows the guidelines of MEC ordinance 1.095, in which consultation by the student's name is one of the ways to validate the diploma. Only the basic information of the diploma is presented in the validation, maintaining privacy for the most personal data. Another situation is diploma revocation, for example due to a plagiarized work. Revocation of a diploma was not addressed in this work, but as a future work this situation can be dealt with in the creation of a status field in the creation of the smart contract, in which it could be changed to active or revoked to control active diplomas and revoked. The use of blockchain has its challenges. When evaluating the blockchain solution practicality, professionals need to carefully evaluate the solutions feasibility that can meet different business requirements.

The benefits of this solution are: helping third parties to verify the records, distributed data in several places, reducing the chance of losing information and having high availability for consultation, since a network node falling or ceasing to exist does not prevent the information from being available for access.

As future work, we intend to conduct a case study on Ethereum and measure application's performance, and look at the issue of transaction costs. It is intended to evaluate the proposal with other educational institutions with the presence of some MEC employee to represent the Brazilian government during the validation. Finally, we plan to extend Educ-Dapp application features and analyze how to make the application more adherent to the criteria of [8].

REFERENCES

[1] K. Al Harthy, F. Al Shuhaimi, and K. K. Juma Al Ismaily. 2019. The upcoming Blockchain adoption in Higher-education: requirements and process. In *2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)*. 1–5.

[2] Roman Beck, Michel Avital, Matti Rossi, and Jason Bennett Thatcher. 2017. Blockchain Technology in Business and Information Systems Research. *Business & Information Systems Engineering* 59, 6 (01 Dec 2017), 381–384. https://doi.org/10.1007/s12599-017-0505-1

[3] N. D. Bhaskar and D. L. K. Chuen. 2015. Bitcoin Mining Technology. In *Handbook of Digital Currency*, David Lee Kuo Chuen (Ed.). Academic Press, San Diego.

[4] Amiangshu Bosu, Anindya Iqbal, Rifat Shahriyar, and Partha Chakraborty. 2019. Understanding the motivations, challenges and needs of blockchain software developers: A survey. *Empirical Software Engineering* 24, 4 (2019), 2636–2673.

[5] Guang Chen, Bing Xu, Manli Lu, and Nian-Shing Chen. 2018. Exploring blockchain technology and its potential applications for education. *Smart Learning Environments* 5, 1 (03 Jan 2018).

[6] J. Cheng, N. Lee, C. Chi, and Y. Chen. 2018. Blockchain and smart contract for digital certificate. In *2018 IEEE International Conference on Applied System Invention (ICASI)*. 1046–1051.

[7] K. Christidis and M. Devetsikiotis. 2016. Blockchains and Smart Contracts for the Internet of Things. *IEEE Access* 4 (2016), 2292–2303.

[8] Claudio Di Ciccio, Giovanni Meroni, and Pierluigi Plebani. 2020. Business Process Monitoring on Blockchains: Potentials and Challenges. In *Enterprise, Business-Process and Information Systems Modeling*, Selmin Nurcan, Iris Reinhartz-Berger, Pnina Soffer, and Jelena Zdravkovic (Eds.). Springer International Publishing, Cham, 36–51.

[9] E. Durant and A. Trachy. 2017. Digital Diploma debuts at MIT - MIT News. http://news.mit.edu/2017/mit-debuts-secure-digital-diploma-using-bitcoin-blockchain-technology-1017. Online; accessed Apr-2020.

[10] A. Grech and A. F. Camilleri. 2017. Blockchain in Education. JRC Science for Policy Report. https://goo.gl/gG4F1w. Online; accessed Apr-2020.

[11] Naghmeh Ivaki, Nuno Laranjeiro, and Filipe Araujo. 2018. A survey on reliable distributed communication. *Journal of Systems and Software* 137 (2018), 713–732. https://doi.org/10.1016/j.jss.2017.03.028

[12] T. Kanan, A. T. Obaidat, and M. Al-Lahham. 2019. SmartCert BlockChain Imperative for Educational Certificates. In *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*.

[13] Kari Korpela, Jukka Hallikas, and Tomi Dahlberg. 2017. Digital Supply Chain Transformation toward Blockchain Integration. In *50th Hawaii International Conference on System Sciences (HICSS)*. https://doi.org/10.24251/HICSS.2017.506

[14] Q. Lu, X. Xu, Y. Liu, and W. Zhang. 2018. Design Pattern as a Service for Blockchain Applications. In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*. 128–135. https://doi.org/10.1109/ICDMW.2018.00025

[15] MEC. 201. Ordinance No. 1,095 - Brazilian Association of Higher Education Maintainers, Brazilian Ministry of Education. http://abmes.org.br/arquivos/legislacoes/Port-MEC-1095-2018-10-25.pdf. Online; accessed Apr-2020.

[16] MEC. 2019. Brazilian Ministry of Education - Higher Education Census. http://portal.mec.gov.br/component/content/article?id=80481. Online; accessed Apr-2020.

[17] Asger B. Pedersen, Marten Risius, and Roman Beck. 2019. A Ten-Step Decision Path to Determine When to Use Blockchain Technologies. 18, 2 (2019). https://doi.org/10.17705/2msqe.00010

[18] N. Rifi, E. Rachkidi, N. Agoulmine, and N. C. Taher. 2017. Towards using blockchain technology for eHealth data access management. In *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)*. 1–4.

[19] R. Taufiq, A. Trisetyarso, Meyliana, R. Kosala, B. Ranti, S. Supangkat, and E. Abdurachman. 2019. Robust Crypto-Governance Graduate Document Storage and Fraud Avoidance Certificate in Indonesian Private University. In *2019 International Conference on Information Management and Technology (ICIMTech)*, Vol. 1.

[20] P. Thakkar, S. Nathan, and B. Viswanathan. 2018. Performance Benchmarking and Optimizing Hyperledger Fabric Blockchain Platform. In *2018 IEEE 26th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*. 264–276.

[21] M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić. 2018. EduCTX: A Blockchain-Based Higher Education Credit Platform. *IEEE Access* 6 (2018).

[22] Z. Wan, X. Xia, and A. E. Hassan. 2019. What is Discussed about Blockchain? A Case Study on the Use of Balanced LDA and the Reference Architecture of a Domain to Capture Online Discussions about Blockchain platforms across the Stack Exchange Communities. *IEEE Transactions on Software Engineering* (2019), 1–1. https://doi.org/10.1109/TSE.2019.2921343

[23] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in software engineering*. Springer Science & Business Media, Springer Heidelberg New York Dordrecht London.

[24] Gavin Wood. 2019. Ethereum: A secure decentralised generalised transaction ledger.

[25] Xiwei Xu, Cesare Pautasso, Liming Zhu, Qinghua Lu, and Ingo Weber. 2018. A Pattern Collection for Blockchain-Based Applications. In *Proceedings of the 23rd European Conference on Pattern Languages of Programs* (Irsee, Germany) *(EuroPLoP '18)*. Association for Computing Machinery, New York, NY, USA, Article 3, 20 pages. https://doi.org/10.1145/3282308.3282312

[26] Xiwei Xu, Ingo Weber, and Mark Staples. 2019. *Blockchain in Software Architecture*. Springer International Publishing, Cham, 83–92. https://doi.org/10.1007/978-3-030-03035-3_5