Analyzing a Blockchain Application for the Educational Domain from the Perspective of a Software Ecosystem

Antonio Welligton Abreu[1,2], Emanuel Ferreira Coutinho[1,2], Wagner Bezerra[1,2]
Delano Maia[1,2], Alan Nascimento Gomes[1,2], Italo Santos[3]

[1]Graduate Program in Computer Science (PCOMP)

[2]Federal University of Ceará (UFC) - Quixadá, Ceará, Brazil

[3]Northern Arizona University - Flagstaff, AZ - USA

siwelligton@alu.ufc.br, emanuel.coutinho@ufc.br, wagnerbragacrf@alu.ufc.br

delanomaia@ufc.br, alanng@alu.ufc.br, italo.santos@nau.edu

Abstract. *Many educational institutions have the challenge of dealing with different data types. Some of this data is related to the issuance of student certificates. Software Ecosystem (SECO) refers to a collection of software products with some degree of symbiotic relationship.Blockchain can be used in different domains and its features of transparency and immutability are key concerns that make it an attractive technology. This article presents a discussion on an educational domain application from a SECO viewpoint that uses blockchain resources to treat the students' certificates from higher education institutions. Our findings were: a model of an educational application that uses blockchain and research opportunities related to SECO and blockchain for educational domain.*

1. Introduction

Nowadays, several educational institutions have the challenge of dealing with multiple data types, coming in diverse formats, produced and made available in different systems. These data are generated and used by several functionalities and employees at the most varied levels (e.g., management of academic activities, student evaluations, registrations, performance, finance, accounting). Another challenge is how educational institutions deal with the task of generating and distributing student certificates [Abreu et al. 2020]. Student certificates are demanded to declare that a person has completed a level of study during a job selection in private and public companies. Hence, certificates exist in society as valuable information authenticated by an institution. However, this document may be fraudulent and misused by persons with malicious intent resulting in misuse.

Software Ecosystem (SECO) refers to a collection of software products with some degree of symbiotic relationship [Messerschmitt and Szyperski 2003], consisting of a set of actors acting as a unit that interacts with a distributed market between software and services, along with the relationships between these entities [Jansen et al. 2009]. Such relationships are often supported by a technological platform or a common market and carried out by exchanging information, resources, and artifacts.

Blockchain technology is integrating different domains (e.g. logistics, agriculture, and health) and it can also be applied in education domain. In this scenario, principles of document authentication, transparency, immutability, and trust are key concerns that

make it an attractive solution [Grech and Camilleri 2017]. Blockchain is a block sequence containing a complete transaction log as a public book, maintained by multiple nodes in a network [Bhaskar and Chuen 2015]. Each node contains an identical copy of this ledger, each block being a logical sequence of transactions, which are permanent, transparent, and unchanging records [Thakkar et al. 2018]. 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. These concepts allow integrity of the entire blockchain to the first block.

In general, Higher Education Institutions (HEIs) have a specialized system to maintain student records of completed courses, where data is structured to be accessed only by institution employees or students on applications to this end, having little or no interoperability. These systems' databases are hosted in a datacenter within HEIs, with restricted access [Turkanović et al. 2018]. In HEI, the certificates issuance and availability are critical roles as it is strong evidence that the student has completed the course [Al Harthy et al. 2019]. Thus, it is essential to prevent degree transcription fraud and a place to check the data issued accuracy by the institution. In this context, blockchain emerges as an alternative solution.

In Brazil, the certificate registration process involves documents from students and HEIs. According to MEC, more than 1 million higher education students graduate throughout Brazil each year, and in 2018 about 1.2 million graduate students [MEC 2019]. These documents are essential to ensure the safety and validity of the legal acts to be produced by this process, and certificate registration must be registered in a specific book, stored in a physical or electronic environment, at the responsibility of each institution [Abreu et al. 2020]. 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 certificate registration in the Official Gazette (DOU), paying a monetary amount for the used space, and the responsibility for publishing the information in 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 [MEC 201]. In this scenario, Educ-Dapp emerged as a web application that uses blockchain to handle student certificates registration.

In this context, this work proposes a discussion of Educ-Dapp, an application in the educational domain, from the viewpoint of software ecosystems. For this, the methodology applied to this work consisted of the following steps: (i) description of the Educ-Dapp (e.g., features, architecture); (ii) description of SECO Educ-Dapp; (iii) SSN modeling; and (iv) identification of research opportunities. The main contributions of this work are: (i) SSN model of an educational application that uses blockchain resources; (ii) dissemination of blockchain technology in an applied manner; and (iii) a set of research opportunities related to SECO and blockchain for the educational domain.

The paper is divided in the following sections: Section 2 outlines the Educ-Dapp application; Section 3 describes aspects from Educ-Dapp SECO; Section 4 discusses some research opportunities; and finally, Section 5 presents conclusions and future work.

2. Educ-Dapp

2.1. Features

This scenario was defined for a private higher education institution, where the participating key users are from the institution's degree issuing sector. 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) number of the institution; (iii) name of graduate student; (iv) ID 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, the flow of operations, and the interaction between actors are fictitious.

The login that makes institutions registration is generated at the contract time and passed on to the government. The government registers educational institutions that can use the application. Institutions make a registration to be enabled to register the diplomas, where the application makes a validation verifying if it has the registration informed by the government. The system confirms the registration made by the institution and prevents anyone from registering with the application. The other entities only have permission to consult the certificate data, searching by the student's ID.

2.2. Architecture

Educ-Dapp architecture is divided into three layers: (i) Application layer: responsible for interacting with entities outside the system; (ii) API layer: composed of the technologies required to create blockchain access; and (iii) Blockchain layer: where the educational data and smart contracts reside. Ethereum platform was used in this prototype, being a more generic blockchain with smart contracts that function as scripts stored in the blockchain. The prototype used a public blockchain without permission, allowing any external entity to consult the data stored in the blockchain.

Some technologies used for prototype development were: **IDE Remix:** Online IDE for smart contract development; **Solidity:** Ethereum's 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.

Metamask plugin was used for smart contract implementation¿ It creates a cryptocurrency wallet to interact with the application, and any transaction that inserts data into the blockchain has to be approved by Metamask. It acts as an intermediary to perform transactions on the blockchain network. After successful registration, users can login to start the degree registration process (include, revoke and consult). 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, managing all transactions through this account. After implementing the smart contract, the front-end application (Educ-Dapp) integration with the blockchain layer was developed.

3. Educ-Dapp SECO

Figure 1 shows a socio-technical network for Educ-Dapp. Around the central platform, various elements collaborate for development, usage, and evolution. Four groups stand out: (i) technologies for application development, (ii) blockchain technologies, (iii) application execution, and (iv) user community. We can identify some integration points with third-party products and suppliers through the network.

Technologies for application development may vary depending on the addition of new features, portability to other development platforms or programming languages. At Educ-Dapp, the blockchain technology used was Ethereum, and several specific tools to support its development were used. To run the application, browsers are necessary and are a common tool and independent of operating systems. Metamask is an essential integration point in this solution, as the browser must have it installed and configured to be a cryptocurrency wallet. Finally, user profiles are specific to the educational domain and can expand to other institutions that work with certificates. Educ-Dapp SECO promotes the integration between technology providers of many types (e.g., infrastructure, database, frameworks, and more) for the use of the central platform, which is the web application by users, focusing on the treatment of student certificates.

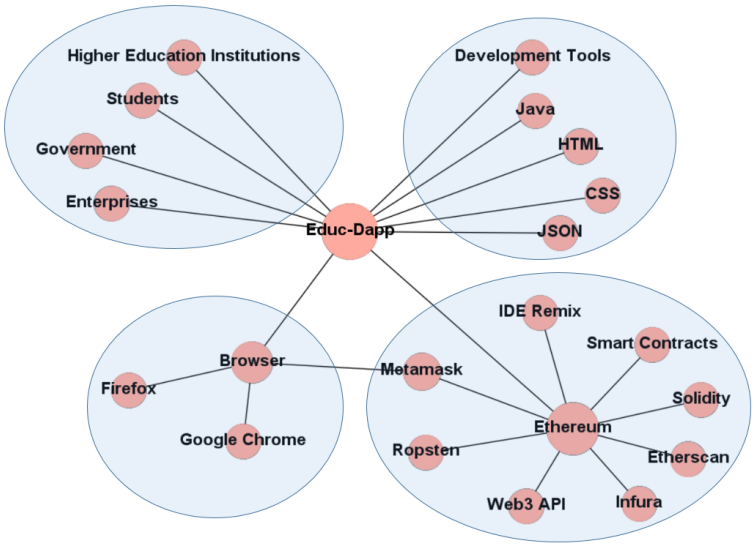

Figura 1. Educ-Dapp sociotechnical network

3.1. Educ-Dapp SSN Modeling

To illustrate the SECO Educ-Dapp, a model designed in SSN (Software Supply Network) was developed (Figure 2). As a company of interest, we have the central platform Educ-Dapp, the educational application where the other elements relate and depend on it. Suppliers are of various types and may undergo modifications throughout the life cycle of the central platform, varying from the application server, programming languages, APIs, IDEs, infrastructure, and utilities. Clients are grouped into companies, higher education

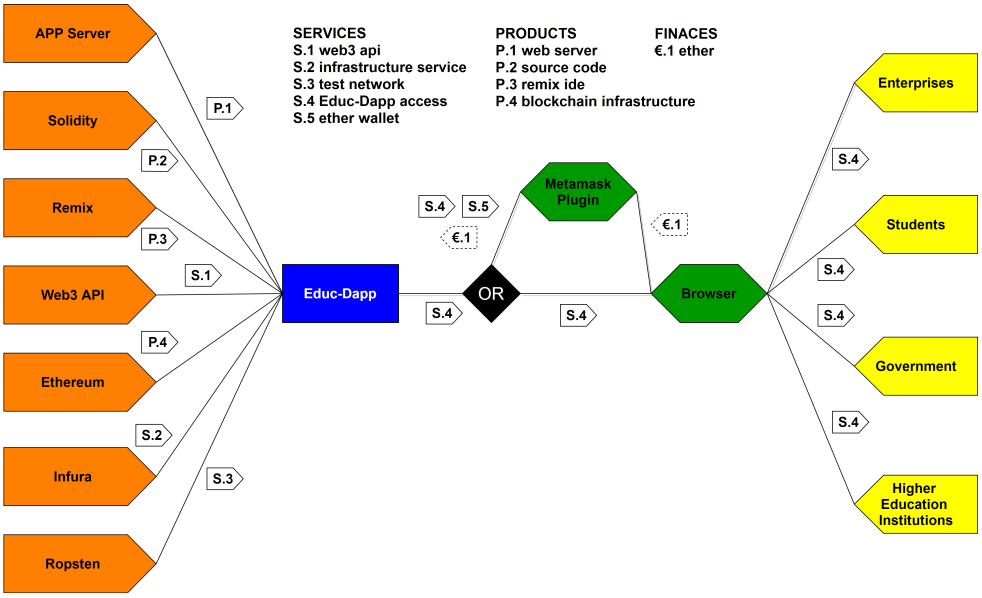

Figura 2. SSN model of Educ-Dapp application

institutions, students, and government intermediaries. We have the metamask plugin and browsers used to access data and applications by users. The relationships between the elements are grouped into services (e.g., API, network, and infrastructure services and facilities for accessing Educ-Dapp), products (e.g., source and executable code and the blockchain database), and finance (e.g., cryptocurrency consumed in operations of blockchain inclusion). In the case of the application's cryptocurrency, which used the Ethereum blockchain, its ether. An SSN component that did not enter the model because there is no possibility of expanding the model was Customer's customer. This point is a possible extension of the model considering that the developer community has access to the Educ-Dapp source code and can develop new solutions and customizations for its customers.

4. Research Opportunities

4.1. Integration between Different Applications

In a SECO, the central platform is often ported to different platforms, such as web and mobile. It implies different versions of applications with diverse technologies and suppliers. Blockchain has several technologies for development and production, such as Ethereum and Hyperledger. How these different versions of applications communicate with the blockchain can vary due to data format, types of requests, and volume of data and access. Permissioned or permissionless [Pedersen et al. 2019] blockchains can impact architecture and integrations. The blockchain itself is not as impacted by different applications, but how it will be accessed can vary. For example, in the educational domain, the Educ-Dapp application is the web and could be ported to a mobile application (e.g., Android, iOS). They are different versions of the same application but with different development platforms. This impacts the maintenance and evolution of the system. Besides, these environments communicate with other applications, which support or complement the central platform, such as the several systems that make up the educational domain. There is a strong characteristic of data integration and sharing, which must be considered for the full functioning of the application.

4.2. Modeling Systems with Different Technologies

Like all systems modeling, SECO modeling must define the levels of abstraction and evaluate the cost-benefit of more or fewer details. A challenge already identified in the literature is the fact that SECO modeling is still very free, varying the level of abstraction [Coutinho et al. 2017], even with the SSN notation used. Static and dynamic aspects of blockchain modeling have already been studied in the literature [Abreu and Coutinho 2020]. For the educational domain, with a strong data integration characteristic, the use of blockchain requires careful modeling, as there are many technologies involved, each with particular characteristics that must be considered in the application design. In this sense, modeling the architecture of educational applications that commonly integrate with several other applications (e.g., financial, academic management, issuing certificates, assessment) can benefit from a SECO model, as the global and relationship view can minimize architectural problems and dependencies.

4.3. Maintenance and Evolution of the Central Platform

Due to technological changes, many maintenance and evolutions can occur on a central platform. Changes in functionality, such as new user profiles that have unique needs or simply changes in scope, impact the central platform. SECO enables the vision of customers, users, and suppliers, which impacts the evolution and maintenance of the central platform. For the educational domain, each institution can have its certificate access application, which accesses a blockchain base. This application can have web and mobile versions and customized features for different educational institutions, which must be considered when maintaining the systems. Blockchain may not impact as much from the functionality perspective. However, it can have a high impact if the structure stored in the blocks varies, which directly affects the database and the applications.

4.4. SECO Testing Blockchain Applications

Software testing research is a robust field with a lot of research that aims to improve testing activities, and there are several related elements in the testing activity that can be further investigated through the concepts defined in the software ecosystem research [Santos et al. 2020]. Blockchain emerges with a vast potential to be applied in different systems and domains, using a specific environment in which it works that no doubt should be intensively tested to guarantee that the blockchain application is behaving as expected. Testing blockchain applications are still in the early stages of research due to the novelty of the technology. However, we believe the importance of testing a blockchain application as Educ-Dapp. It would be necessary to understand how it was previously implemented, therefore understanding a blockchain architecture is an important stage for us to plan and execute future tests. A research opportunity that can be investigated in blockchain applications is to execute performance tests to verify how the blockchain behaves in terms of responsiveness and stability in a specific workload if it can expand and support more users' access. Another important aspect is to perform security testing in a blockchain to guarantee that the data are protected among the users' communication and ensure that the blockchain is free from any possible threats that can cause data loss.

4.5. Social Aspects

A blockchain application involves different stakeholders and can impact users' lives on different levels. These impacts should be further investigated. Understanding how a

blockchain can directly or indirectly impact users, it is possible to know how blockchain systems can be used to help users' activities. In our scenario, Educ-Dapp impacts students who need access to their certification, also the HEIs involve public staff to interact with the system and provide the correct information for the end-users. Another social aspect involved by blockchain technology is how companies (public or private) use it, what are the advantages and disadvantages that come from that context and impacts their users, how we as researchers can develop blockchain systems to maximize the benefits and improve users work environment to make easier to perform their tasks.

4.6. Economic Aspects

An application with a distributed database available to society brings several benefits. However, some economic challenges also arise. One of these challenges is the cost-benefit analysis of migrating production applications that work to new blockchain infrastructure. Many educational institutions have their infrastructures, and when using blockchain, either build new infrastructure or use a third party base. Data access must also be designed to be free or paid. This depends on the business model to be used and impacts end-users. The learning curve for developing blockchain applications should also be analyzed, as it has a cost. Developers must prepare for this new technology and integration aspects between systems.

4.7. Data Security

Considering the presented scenario of the Educ-Dapp application, some research opportunities and challenges can be listed related to access control and data security of students' diplomas [Abreu et al. 2020]: (i) there are several different databases with their 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 a challenge, as students need the HEI that issued the certificate to prove their training in the face of a studies continuation or job selection; (iii) the fact that it does not depend on a manual administrative process or the availability of an HEI system to prove a certificate can be an improved situation to be considered; (iv) each HEI is responsible for the safe storage of the students' diplomas data, since mechanisms used by each HEI may not be the most adequate to safely store the data for an extended period, causing the loss of information; and (v) preventing forged documents is one of the significant challenges of higher education, and it is necessary to have some data verifiability mechanism to avoid certificate falsification.

5. Conclusion

This study presented an educational application called Educ-Dapp that uses blockchain resources to treat the students' certificates from higher education institutions. Despite the application being a prototype, it is understood that blockchain provides suitable solutions for security, integration, and data traceability, interesting features for the educational domain. The main contributions of this work are: (i) SSN model of an educational application that uses blockchain resources; (ii) dissemination of blockchain technology in an applied manner; and (iii) a set of research opportunities related to SECO and blockchain for the educational domain. As future work, we intend to deepen each presented challenge to collaborate with the blockchain community, specifically in the educational field.

Referências

Abreu, A. W. and Coutinho, E. F. (2020). Motivating web and blockchain application modeling. In *2020 IEEE International Conference on Software Architecture Companion (ICSA-C)*, pages 110–113.

Abreu, A. W. S., Coutinho, E. F., and Bezerra, C. I. M. (2020). A blockchain-based architecture for query and registration of student degree certificates. In *Proceedings of the 14th Brazilian Symposium on Software Components, Architectures, and Reuse*, SBCARS '20.

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

Bhaskar, N. D. and Chuen, D. L. K. (2015). Bitcoin mining technology. In Chuen, D. L. K., editor, *Handbook of Digital Currency*. Academic Press, San Diego.

Coutinho, E. F., Viana, D., and d. Santos, R. P. (2017). An exploratory study on the need for modeling software ecosystems: The case of solar seco. In *2017 IEEE/ACM 9th International Workshop on Modelling in Software Engineering (MiSE)*, pages 47–53.

Grech, A. and Camilleri, A. F. (2017). Blockchain in education. jrc science for policy report. https://goo.gl/gG4F1w. Online; accessed Apr-2020.

Jansen, S., Brinkkemper, S., and Finkelstein, A. (2009). Business network management as a survival strategy: A tale of two software ecosystems. In *Proceedings of the First International Workshop on Software Ecosystems, 11th International Conference on Software Reuse*, volume 505.

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.

MEC (2019). Brazilian ministry of education - higher education census. http://portal.mec.gov.br/component/content/article?id=80481. Accessed Apr-2020.

Messerschmitt, D. and Szyperski, C. (2003). *Software Ecosystem: Understanding an Indispensable Technology and Industry*. The MIT Press, The MIT Press, 1 edition.

Pedersen, A. B., Risius, M., and Beck, R. (2019). A ten-step decision path to determine when to use blockchain technologies. *MIS Quarterly Executive*, 18(2).

Santos, I., Coutinho, E. F., and Souza, S. R. S. (2020). Software testing ecosystems insights and research opportunities. In *34th Brazilian Symposium on Software Engineering*, SBES '20, page 421–426.

Thakkar, P., Nathan, S., and Viswanathan, B. (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)*.

Turkanović, M., Hölbl, M., Košič, K., Heričko, M., and Kamišalić, A. (2018). Eductx: A blockchain-based higher education credit platform. *IEEE Access*, 6.