A Blockchain-based Conceptual Model to Address Educational Certificate Verification Challenges in Tanzania

Said Hamisi Said
School of Computational and Communication Science and Engineering, The Nelson Mandela African Institution of Science and Technology, Tanzania
saids@nm-aist.ac.tz (corresponding author)

Mussa Ally Dida
School of Computational and Communication Science and Engineering, The Nelson Mandela African Institution of Science and Technology, Tanzania
mussa.ally@nm-aist.ac.tz

Efraim Michael Kosia
The Nelson Mandela African Institution of Science and Technology, Tanzania
efraim.kosia@nm-aist.ac.tz

Ramadhani S. Sinde
School of Computational and Communication Science and Engineering, The Nelson Mandela African Institution of Science and Technology, Tanzania
ramadhani.sinde@nm-aist.ac.tz

ABSTRACT

The proliferation of counterfeit educational certificates is an ongoing issue around the world, including Tanzania. The effect of this malpractice is detrimental to the credibility of education. Traditional strategies to prevent fake certificates are abortive, calling for a more sophisticated approach. Blockchain technology has recently emerged as an ideal solution to this problem due to its inherent attributes that ensure disintermediation, immutability, tamper proof, anonymity, transparency, consensus, security, and trust. However, most existing blockchain-based solutions lack crucial functionalities that are pertinent to the Tanzanian education system. This study unveiled the challenges faced by the current verification system in Tanzania and proposed a blockchain-based conceptual model to address them. The proposed model is based on blockchain, smart contracts, and the Interplanetary File System (IPFS). Quantitative and qualitative methods were used to investigate certification problems in Tanzania and modeling techniques were used to construct the conceptual model. The findings showed that the main challenges of the current verification system emanate from manual procedures, unverifiable credentials, susceptibility of centralized storage systems, disintegrated verification systems, revocation problems, difficulties in communication, and high dependency on the issuers. These challenges undermine certificate verification, impose a significant setback in the fight against forgeries, and create loopholes for forgeries to persist. It was conceptually demonstrated that these issues can be resolved through the proposed blockchain-based solution.

Keywords: blockchain; smart contract; interplanetary file system; educational certificates; counterfeiting; verification; Tanzania

I. INTRODUCTION

Educational certificates are documents used to certify an individual's achievement in education and provide proof that the holder possesses the knowledge and skills required to qualify for different responsibilities [1-4]. Possession of such credentials can provide an individual with social and economic opportunities [3, 5]. As these benefits are desirable, some
individuals who cannot endure the cost, time, and hassles required to obtain them resort to cheating on their qualifications by producing counterfeit certificates [5-6]. With the advancement of multimedia and printing technologies, the production of counterfeit certificates has become easier and their quality has improved to such an extent that it is difficult to distinguish between fake and genuine [7]. Counterfeiting of certificates can involve the fabrication of a new certificate (using fake seals, signatures, etc.), the modification of a legitimate certificate (by altering its content such as grades, etc.), degree mills (acquiring certificates from fictitious learning institutions), the in-house fabrication of certificates by corrupt officials of a legitimate institution, a mistranslation of foreign certificates, impersonation, and the deceitful use of a revoked certificate [8-11]. Moreover, it may also involve hacking the issuer's database to modify educational records [12].

The sale of fake certificates has become a lucrative business with a billion-dollar market around the world [9, 13]. According to [14], there are more than 2 million fake degrees in the United States. In [8], the existence of around 810 fake learning institutions in the United States and about 272 in the United Kingdom was reported. According to a report by the Indian authorities, around 40,000 people in Bangalore gained employment using fake academic qualifications [15]. In Tanzania, in 2017, the verification of secondary school and teacher certificates of 400,035 public servants revealed that almost 10,000 possessed fake certificates [16]. The effect of this malpractice not only damages the reputation and credibility of education but also leads to poor performance and productivity at work and deprives genuine graduates of their deserved opportunities [4-5, 17]. Prevention of this menace is mainly based on verification strategies [10, 18]. Several verification strategies against certificate forgery have been proposed and established, but the problem persists, showing that the currently employed methods are not able to tackle it and there is a need for more sophisticated approaches. Blockchain technology has recently emerged as a potential tool to solve problems that seemed impossible using traditional technologies, due to its ability to build trust-free, transparent, and distributed systems that run on a peer-to-peer consensus network through which strangers with different interests can transact without relying on intermediaries or central authorities for trust maintenance [10, 19-20]. This technology was brought to light in 2008 through the Bitcoin cryptocurrency [21]. Since then, it has been enormously transformed and its application has gained ground in various domains [22-24], including document verification for fraud detection [2, 25]. Blockchain can substantially enhance credential verification and effectively combat counterfeiting [5, 23] due to its innate properties that ensure: disintermediation of management, immutability and tamper-proof of records, anonymity of users, consensus and transparency of transactions, security of the system, and trust between parties [1, 23, 25].

However, to fully realize the potential of blockchain-based certification solutions, their implementation should be tailored to the contextual needs of a particular education system, as the education system differs from country to country [26]. This study investigated the challenges of the traditional verification system in Tanzania and proposed a blockchain-based conceptual model tailored to them, demonstrating how it can address them. The proposed blockchain-based conceptual model is unique in the sense that it can address the certification problem at all national education levels (from secondary schools to university), incorporates regulatory authorities to manage registration and deregistration of issuing institutions to the system, controls access permissions of different roles involved in the system, allows issuance and on-chain revocation of certificates, allows third-parties to verify certificates by directly querying the blockchain, and uses decentralized off-chain storage alongside blockchain to facilitate storage of raw data. Finally, it facilitates the issuance, revocation, and verification of the “equivalent statements” for graduates who acquired certificates abroad.

II. BACKGROUND

A. Certification and Verification Ecosystem in Tanzania

The educational qualification verification ecosystem consists of four major entities, regulators, issuers, graduates, and recruiters, who work together to achieve the verification life cycle, as shown in Figure 1.

![Fig. 1. Educational qualifications verification ecosystem in Tanzania.](Image)

The government, through the respective ministry of education, establishes the regulatory bodies for different types and levels of education in the country. These regulators include the Tanzania Commission for Universities (TCU) which is responsible for universities [27], the National Council for Technical and Vocational Education and Training (NACTVET) for technical and vocational institutions [28], the National Examination Council of Tanzania (NECTA) that is mandated to oversee national examinations [29], and the Ministry of Education Science and Technology (MoEST) as the general custodian of education [30]. These bodies are responsible for, among other things, approving the issuers of...
educational qualifications in the country. As such, they publish a list of approved issuers to which third parties can refer if they wish to ascertain the legitimacy of the issuing institutions. According to Tanzanian regulations, they also evaluate the educational qualifications awarded outside the country, determine their comparability with the Tanzanian qualifications standard, and issue a comparable certificate called the educational qualifications awarded outside the country. According to Tanzanian regulations, they also evaluate the educational qualifications in the country. As such, they publish a mistake or if the graduate commits academic or non-academic misconduct. After receiving a certificate, graduates can present it to prospective recruiters during job applications, admissions, or other purposes. Recruiters must verify the validity of the certificates presented to ensure that they do not enlist candidates with forged credentials. However, they cannot achieve this task without the support of the issuers and regulators. Therefore, in addition to performing internal verification, recruiters usually consult with issuers to verify certificates issued by institutions within the country. In case of certificates granted outside the country, the recruiters consult the regulators, who are responsible for assessing and issuing equivalent statements for foreign awards. Recruiters must also confirm that the learning institution exists legally. To do so, they consult a list of approved issuers published by the respective regulators on their websites.

B. Blockchain Technology

Blockchain embeds many technologies, including cryptography, hash functions, peer-to-peer (P2P) networks, and consensus algorithms [3, 32]. Blockchain can be defined as a distributed digital ledger consisting of a growing chain of transactional records sealed in blocks that are linked together using a cryptographic hash to ensure security [25, 33-34]. Through their nodes, users of the blockchain constantly add new cryptographically signed transactions to the ledger. Special users, called miners, seal the transaction into a block and attach it to the blockchain ledger [35]. Before appending, the blocks are hashed using a one-way hash function, such as the Secure Hash Algorithm (SHA)-2 or 3, and then chronologically linked by including a hash value of a previous block in a newly added block [20, 36] to form a chain of blocks. This mechanism ensures the integrity, immutability, and tamper-proof of the data in the ledger [3]. Figure 2 shows the structure of a block and the way they are linked. A block consists of a header and content. The content holds transactions, while the header contains metadata such as block hashes (for current and previous blocks), a timestamp (block creation time), and a random number (nonce) to verify the block’s hash [33, 37]. The ledger and its data are then replicated among the nodes in a P2P network, making it distributed and decentralized [35, 38-39]. Being decentralized and distributed means that no single node acts as a central server, as each node holds data, and the transaction of each node is visible to all other nodes, making the blockchain transparent, secure, reliable, and trustworthy [3, 38, 39]. Any update to the ledger can only occur if most nodes agree to the change to ensure the integrity, transparency, and authenticity of the transactions [2]. The consensus mechanism is used to reach an agreement among the nodes on the node to add a new block and on the block to be added to the blockchain [40]. This consensus is handled by algorithms such as Proof-of-Work (PoW), Proof-of-Stake (PoS), Delegated Proof-of-Stake (DPoS), and Byzantine Fault Tolerance (BFT), depending on the blockchain platform [38, 41]. So far, there are many blockchain platforms, but the most popular are Bitcoin [21], Ethereum [42], and Hyperledger [43]. Blockchain can be public, private, or consortium, and its access can be permissionless or permissioned [36, 38]. A public blockchain, sometimes called permissionless, is free for anyone to join the network, read or write to the ledger, validate transactions, and mine the blocks [44, 45]. A private blockchain, also called permissioned, is owned by a single organization that is responsible for managing the blockchain and controlling access, and reading and writing the ledger [45]. Unlike private and public, a consortium blockchain is owned by a group of organizations to provide an inter-organization service. In this blockchain, the consensus process is carried out by participating organizations, while reading or writing permissions may be public, or restricted to the consortium members [40, 46].

Fig. 2. The blockchain data structure, showing block structure and interconnections.

With blockchain, strangers can transact without relying on intermediaries for trust keeping [10]. Instead, trust is ensured by asymmetric cryptographic algorithms [19], such as RSA, which is a prominent algorithm in asymmetric cryptosystems [53]. These attributes made blockchain a disruptive technology with broad applications in various domains such as finance [21], health [47], agriculture, and education [48].

C. Smart Contracts

In its first generation, blockchain was only associated with cryptocurrency, but in the second generation, smart contracts were introduced to make it programmable and applicable beyond cryptocurrencies [23, 49]. Smart contracts are computer programs (codes) stored on the blockchain that run automatically when the required conditions are met [33, 35, 50-51]. The concept of smart contracts, though recently
popularized in blockchain applications, was first introduced in 1997 [52]. Typically, all blockchains contain smart contracts as built-in scripts to execute their transaction logic [34]. However, most blockchains, such as Bitcoin, offer very limited programmability. Ethereum blockchain is a pioneer in supporting Turing-Complete smart contracts that can perform general-purpose programming [32, 35, 53], as it provides a dedicated high-level programming language and an execution environment named Ethereum Virtual Machine (EVM) [32-33, 35, 53-54]. In Ethereum, smart contracts are created by nodes through a special transaction, and they are assigned a unique address on the blockchain [55]. Similarly, they are executed only when triggered by a transaction directed at them through their address, specifying the input data and the function to be called [37, 55]. They are executed independently and exactly as programmed without the possibility of censorship, falsification, or interference by third parties [46]. The introduction of smart contracts has enabled blockchain to handle more complex business constraints and promoted its application to domains other than cryptocurrency. The core functionalities of the proposed blockchain-based conceptual model in this study, such as the issuance, revocation, and verification of certificates, are demonstrated using smart contracts. Smart contracts are also used to manage the registration, deregistration, and access control of different roles involved in the model.

D. InterPlanetary File System (IPFS)

Although blockchain provides useful properties, its computational power and storage space remain limited and expensive [1, 35, 37]. For that reason, it is common practice to store raw data off-chain while storing only a few critical data, such as hashes and meta-data on the chain [35]. To address this need, several off-chain data storage solutions have been developed, including IPFS, Swarm, Sia, and StorJ [56]. Of all, IPFS [57] is the most popular and widely used. IPFS is a decentralized and secure content-addressable file storage system that distributes stored contents among nodes in a P2P network [33, 58]. IPFS is content addressable in the sense that each content (file), locally stored in a peer node, is hashed and the generated hash is used as a content identifier (CID) [33, 58]. During retrieval, the content is located by specifying its CID rather than its location [58]. The proposed model uses IPFS for the storage of raw certificates, equivalent statements, issuer profile information, and other complementary information to relieve the blockchain from storage overhead.

III. RELATED WORKS

A. Non-Blockchain-based Related Works

Before the advent of blockchain, scholars and practitioners used non-blockchain-based techniques to improve the verification of educational certificates and alleviate associated fraud [59]. These methods came in a variety of ways and used different techniques, such as OCR [60], hash [61], 2D barcode [62], QR codes [63], RFID [64], NFC [65], web applications [7], and mobile applications [66]. In [13], poor verification methods were shown to fail to detect forgeries. Furthermore, since the verification process left loopholes for bad actors to use fake certificates, the digitization of the verification process was suggested at all levels of education. In [66], QR codes and web and mobile applications were used to develop a certificate verification system for Semarang University in Indonesia. This approach formulated a fingerprint of the certificate information, which was encoded along with a URL link to the web application as a QR code. For verification, the verifiers scan a QR code using their smartphone to redirect to the web app where information is stored for comparison. In [7, 67], challenges in verifying educational credentials against counterfeits were found, and web-based applications were developed to authenticate credentials. These approaches were somewhat similar, as they both involved centralized databases of student records and provided an Application Programming Interface (API) for recruiters (e.g. employers) to access them via a web-based application. However, these techniques are manual or semi-automated, making them inefficient for certificate verification and forgery prevention [59]. Furthermore, most of them depend on conventional centralized database systems. As a result, they are vulnerable to a single point of failure [10, 26, 56], internal [4, 68] and other security threats.

B. Blockchain-based Related Works

Several initiatives and research efforts have been presented on utilizing blockchain for educational certificate verification and forgery prevention. Blockcerts [69] is one of the early blockchain-based initiatives on certification. This is an open standard platform based on Bitcoin that enables educational institutions to implement blockchain-based solutions for issuing, sharing, and verifying certificates. However, as noted in [70], it cannot verify the issuing institutions, as it does not check if the public key is owned by the issuer, leaving a chance for fake institutions to issue certificates. Moreover, it does not implement on-chain certificate revocation [71]. In 2017, the University of Nicosia (UNIC) used Blockcerts and became the first institution to issue all academic certificates on the Bitcoin blockchain [72]. In their solution, a PDF certificate is hashed using SHA-256, and the digest is stored on the blockchain so that third parties can access it to verify the certificate. This solution is limited to a single university and inherits Blockcerts limitations. Capitalizing on the same Blockcerts, Malta is said to be the first country to issue educational certificates on the blockchain at a national level [73]. This solution is said to operate at the national level, but its technical details have not been presented. However, the use of Blockcerts may embed its limitations.

Many other solutions followed with some improvements. For example, using the ARK blockchain, EduCTX was proposed in [36] to improve the issuance, storage, sharing and verification of credits accrued from higher education. The system is based on the European Credit Transfer and Accumulation System (ECTS) and represents ECTS credits as ECTX tokens on the blockchain. In [1], Blockchain for Education was proposed using the public Ethereum blockchain, smart contracts, IPFS, and proof-of-stake to issue, share, and validate certificates. In [26], the University of Zurich BlockChain (UZHBC) was proposed for certificate verification, using the public Ethereum blockchain and smart contracts. In [10], a comprehensive blockchain-based solution, called Cerberus, was presented for the verification of higher
IV. METHODOLOGY

A. Uncovering the Certification Problems

1) Research Approach and Design

Quantitative and qualitative research approaches were used to provide a deeper and broader understanding of the certification problem. The case study design strategy was adopted, as suggested in [75-76], to thoroughly investigate the challenges of certificate verification in the case of Tanzania.

2) Research Setting and Participants

This study was conducted in the Tanzanian context because it is intended to uncover and address the challenges of certificate verification in Tanzania. The study focused on secondary, vocational, technical, and university certificates, which are the most used in applications for jobs, admissions, and other purposes. Information on certification issues was gathered from certificate issuers (educational institutions), recruiters (employers and admission officers), and regulators of educational institutions.

3) Sample Size and Sampling Technique

A total of 137 participants were involved, including 65 issuers, 67 certificates, and 5 regulators. Cochran's formula for the unknown population and its adjustment formula for the finite population were used to determine the sample size for issuers and recruiters, whereas, for regulators, a complete enumeration was used [77-78]. The respondents were selected using a nonprobabilistic sampling technique called Stratified Purposive Sampling [79-80] to focus on cases that are rich in desired information and can accommodate the heterogeneity of the research population (certification stakeholders), which is divided into different groups [81].

4) Data Collection and Analysis

This study adopted methodological and data source triangulation techniques in data collection to ensure the validity and reliability of its findings, as suggested by [82-83]. Data were collected from multiple certification stakeholders using questionnaires and interviews, supplemented by observations and documentary reviews to capture different perspectives on the certification problem in Tanzania. For data analysis, descriptive and thematic approaches were used to analyze quantitative and qualitative data, respectively.

B. Designing the Conceptual Model

The design concept of the model relies on existing knowledge from the literature and other available resources. Its construction involved the use of design modeling techniques and tools to represent and describe the structure, components, relationships, behavior, and functionality of the proposed artifact. The Unifier Modeling Language (UML) [84] was used as a general guideline to inform the design, and the diagrams.net [85] diagramming tool was used for creating a visual representation of the model.

V. RESULTS AND DISCUSSION

A. Current Verification Process and Methods

The verification process begins when recruiters receive certificates from their candidates as they submit applications via email, hand delivery, or online application systems. After receiving a certificate, the recruiters verify it using several methods, as shown in Figure 3. From the findings, 76.1% of the recruiters reach out to issuers for verification, 71.6% ask their applicants to get their certificates certified by a notary or issuer, 64.2% inspect the certificate's security features, 43.3% indicate that, apart from checking the certificates, they also examine the competence of the candidates via aptitude tests, 29.9% verify online via an API that links their online application system to issuer's centralized database, however, this is only used for secondary certificates and only applicable to a few recruiters who use online application systems, 23.9% accept a certificate by a mere trust or its look without establishing concrete facts, and 4.5% verify by scanning QR codes imprinted on the certificates. The results show that the most widely used method is reaching out to issuers for verification, implying a high dependency on the issuers for verification. This dependency is attributed to the fact that the proof-of-authenticity features embedded in the certificates can only be confirmed by the issuers, and issuers are the only ones who maintain student records that can be referred to for verification of certificates. QR code scanning is the least used method because it is an emerging technology in Tanzania and only a few institutions use such codes.

eduction certificates, combining the Ethereum blockchain with QR codes to improve usability and using on-chain smart contracts to improve certificate revocation. Other similar solutions, such as SmartCert [74], VECefblock [38], and Edu-Dapp [37], were also proposed to overcome the problem of fake credentials during recruitment.

In general, these solutions were oriented to the requirements of particular educational systems. For example, VECefblock was designed following the Vietnamese education structure, UZHBC is specific for the University of Zurich, UNIC's solution is restricted to their university, and EduCTX is meant for higher education institutions that follow ECTS standards. For that reason, they cannot directly fit into solving the certification problem in the Tanzanian education system. As suggested in [26], to fully realize the potential of blockchain-based certification solutions, its implementation should be tailored to the contextual needs of a particular education system, because the education system differs from one country to another. Therefore, the proposal of this study is tailored to the setup of the Tanzanian educational certification system to address the contextual challenges correctly. Moreover, apart from issuance and verification of certificates, it incorporates regulatory authorities to manage issuing institutions, controls the role's access permissions, provides on-chain certificate revocation, uses decentralized off-chain storage, and takes into account the equivalent statements for certificates acquired abroad. The existing solutions have not included all these features, hence partially solving the certification problem.
As mentioned above, in addition to conducting internal verification, the recruiters also consult issuers to verify certificates. Figure 4 shows that after receiving verification requests, the issuers verify using the following methods: All issuers (100%) indicated that they consult student or certificate records to match the information, 73.8% indicated that they inspect the certificate's security features, and 9.2% indicated that they scan QR codes embedded on the certificates. However, most of them use a combination of methods. The reasons that contribute to the low adoption of QR codes by recruiters are also applicable to issuers.

2) Unverifiability of Certificates

Approximately 70% of the respondents indicated that unverifiable credentials from recruiters are one of the challenges facing the verification of certificates. To protect a certificate from forgery, issuers usually incorporate physical security features such as holograms, signatures, stamps, and others to act as a tamper-proof and proof-of-authenticity. However, these features are secretly encoded by the issuers to avoid imitation, and they are the only ones who can decode them. Consequently, recruiters cannot verify the authenticity of certificates because they are unfamiliar with the features.

3) Communication Drawbacks

To reliably verify the certificate, recruiters have to seek assistance from issuers and regulators. They communicate by sending letters with certificates to be verified, which are then delivered by physical visits, postal mail, or email, while sometimes they also use phone calls to inquire for verification. As indicated by 59.9% of the respondents, current communication methods pose some serious challenges that impede the verification of certificates for fraud prevention. These methods are associated with delayed delivery of verification requests and responses. As a result, a lot of time is spent either reaching out to the institutions or waiting for verification feedback. Sometimes, the request may not be responded to. Additionally, there is a significant financial burden in facilitating physical visits and communication. For instance, an employer responded that "contacting issuers to verify certificates takes a very long time, like weeks or even months, and, worse enough, they may not respond to our requests. Physically visiting each issuing institution is not feasible because it is more expensive and time-consuming."

4) Vulnerability of Centralized Systems

The issuing institutions maintain graduate records in their registries, stored electronically in centralized computer systems along with physical storage in file cabinets. Issuers retrieve the information required from these records to print or verify certificates. Therefore, the security of these systems is crucial. However, since these systems are centralized, they are vulnerable to attacks and difficult to secure, as indicated by 40.1% of the respondents. As they act as a single point of failure, both issuers and recruiters reported several instances in which they were unable to perform verification due to system failure or unavailability. Second, they are susceptible to internal threats posed by corrupt officials, who may collude
with graduates to manipulate records and print fake certificates. Most issuers admit that it is hard to control internal system users, especially super users, from tampering with records.

5) Certificate Revocation Issues

Approximately 36% of the respondents expressed that an inappropriate revocation mechanism imposes significant challenges in the verification of certificates. As mentioned above, issuers can revoke a graduate certificate in case of a problem. The current revocation mechanism involves writing a letter to the graduate and issuing a public notice of the annulment. However, this approach does not provide a way to retract a certificate from the revoked person. That means that the person may continue to use it illegally despite the revocation. Public notice alone, as is currently done, is not enough, because it cannot be readily available to recruiters whenever they need to verify a certificate. As a result, most of the recruiters do not check whether a presented certificate is revoked or not, as asserted by one of the recruiters: "We have no way to tell whether the certificate is revoked or not."

6) Disintegrated Verification Systems

During the recruitment application, the candidates must submit several certificates acquired from different educational institutions. To verify the credentials of a single candidate, recruiters need to check with each institution, as there is no single point through which all the certificates can be verified. The process of consulting many institutions for verification is described as tedious, time-consuming, expensive, and inefficient, especially when having many candidates to verify. Approximately 72% of the respondents pointed out this issue as one of the bottlenecks of the verification process. This fact was reinforced by one of the recruiters, who stated that "there is no common repository of graduates' information that includes data from all issuers, which recruiters could access through a single interface for verifying all the certificates. As a result, the verification of certificates becomes difficult."

7) Dependency on Issuers for Verification

As indicated by 76.1% of recruiters, reaching out to issuers is the most used method for verifying certificates, implying a high dependency on the issuers for verification. This overdependency on the issuers imposes some challenges to both issuers and recruiters, as indicated by 56.2% of the respondents. On the issuer side, they are overloaded by the heavy workload imposed by a large volume of verification requests they receive from recruiters across the country. On the other hand, recruiters spend a lot of time and money reaching out to issuers for verification while pending the recruitment to wait for the verification outcome.

C. The Proposed Blockchain-based Conceptual Model

1) Architecture Overview

This study proposed a model to solve the identified certificate verification challenges in Tanzania. Figure 6 shows the proposed model based on blockchain, smart contracts, IPFS, and a decentralized application.
This model supports the issuance, revocation, and verification of certificates and equivalent statements on the blockchain. Since the proposed model aims to improve certificate verification and forgery prevention during recruitment, it focuses on certificates that are applicable during recruitment (i.e. secondary, vocational, technical, and university certificates). This model involves the same actors as in the traditional verification system, that is, a central authority responsible for establishing regulatory bodies, regulatory bodies responsible for approving issuing institutions, issuing institutions, graduates, and recruiters (employers, admission officers).

This architecture allows users to join the blockchain network through their respective front-ends on the decentralized application (DApp). The DApp intermediates the interaction between users, the blockchain, and IPFS. The DApp is powered by smart contracts that execute on the blockchain. Smart contracts are used to implement the core logic of the model, contain variables to store certificate hashes and their associated metadata, and functions and methods for handling all the operation logic over them. Contracts can be implemented using programming languages, such as Solidity, and deployed in any blockchain platform that supports smart contracts, such as Ethereum. IPFS is used as decentralized off-chain storage of files and their associated metadata. Due to the storage and computational limitations of blockchain, if everything is performed on-chain, its performance becomes poor. Therefore, shifting some of the data and activities off-chain will potentially boost the overall performance.

2) Management of User Roles and Permissions

The management of user roles and permissions is a crucial aspect of the model. This mechanism ensures that only authorized institutions can join the blockchain to issue or revoke certificates [37], and, in turn, fake institutions can be prevented from issuing certificates [86]. Furthermore, it can also provide a way for the government and its authorities to monitor the issuance of certificates [86]. Smart contracts can be programmed to provide some checks, ensuring that only users with the appropriate permissions can perform transactions in the ledger [32, 46]. Figure 7 shows a hierarchical scheme of roles in the proposed model that involves the central authority, regulators, and issuers. Graduates and recruiters are not included in this hierarchy because they only retrieve information from the ledger, which does not involve any transaction. The central authority (e.g. the Ministry of Education representing the government) will initiate the smart contracts and act as their administrator. Through smart contracts, the central authority will register the regulators (such as TCU, NACTVET, etc.), who in turn will register the issuers. Since there are different issuing institutions at different levels of education and each level has its regulatory body, the model is designed to involve several regulators, and through smart contracts’ configuration, each regulator will be allowed to register its respective issuers to the blockchain. After joining the blockchain, the regulators will register their respective issuers on the blockchain and grant them permission to issue and revoke certificates. According to Tanzania regulations, apart from approving the issuers, the regulators are also mandated to suspend or stop the issuer from issuing certificates if they fail to meet the required quality standards. Using smart contracts, the model also allows regulators to remove issuers from the blockchain when needed. Smart contracts act as the brain behind this model and contain features such as structs, mapping, and functions. The "structs" will be used for storing user information, "mapping" (a key-value data structure) will be used to associate the user's blockchain address with their stored details, and "functions" will be used to implement all the operation logic for user registration, deregistration, and permissions. To facilitate registration, a function can be implemented to accept user registration details (such as name, address, etc.) and associate them with the specific addresses on a smart contract. For deregistration, another function can be implemented to allow user removal by deleting their details from the contracts. Functions can also be implemented to define user roles and restrict access and permissions. Due to storage limitations, the blockchain only stores essential details, such as CIDs, and raw files containing issuer profile information will be stored in the IPFS. Using CIDs, recruiters will be able to retrieve the corresponding files from the IPFS, allowing them to confirm the legitimacy of the issuers during certificate verification.

3) The Issuance Process

To issue certificates (as shown in Figure 8), registered issuers will begin by preparing digital certificates in PDF or other formats that encode the certificate's information, such as the issuer's name, student's name, student's picture, graduation date, certificate identity, award title, and grade [11, 26]. The PDF file format is more appropriate for compatibility with legacy systems because PDF documents from the Student Record Management System (SRMS), which are normally used to print paper-based certificates, may now be used as input to the proposed system. This document and other associated metadata (such as the graduate's blockchain address) are uploaded to the DApp. The DApp forwards it to IPFS, where it is hashed using a deterministic one-way cryptographic hash function to generate a unique hash value. While the raw files are stored in the IPFS, the hash value is returned to the DApp, which then pushes it to the blockchain. To achieve this, the DApp invokes the smart contract function through a transaction signed using the issuer's private key and passes the hash value as an argument [39]. The smart contract function will verify the authenticity of the transaction source, ensuring that it comes from an authorized issuer. The function then stores the hash value alongside the certificate metadata, such as its revocation status and others, in the ledger. After that, a transaction ID/hash is returned to the DApp, then to the issuer, and the issuer shares
it with the graduate along with the digital certificate. The graduate can also obtain the issued credentials on the DApp. Through the DApp, graduates will be able to retrieve, view, and download certificates. Only the hash values of the certificates and a few attributes, such as their revocation status, are stored on-chain. Real digital certificates will be stored off-chain on the IPFS.

In the blockchain, graduates are identified using their blockchain addresses, which are generated using wallets such as MetaMask. Wallets typically generate public and private keys, which are used to create the corresponding blockchain addresses [33]. In the proposed model, graduates will be able to set up their wallet accounts, through which they can generate and manage their cryptographic keys and their corresponding blockchain addresses [10]. The wallet is then connected to the DApp through which the blockchain address can be shared with the issuers to be used to issue a certificate. Each certificate hash stored on the blockchain must be affiliated with the blockchain address of the corresponding graduate. This is achievable by using the "mapping" data structure of smart contracts. To make the model applicable across different levels of education, the same blockchain address will be used at every level of education attended by the graduate to ensure that all the certificates acquired throughout his educational journey are associated with the same address.

4) The Revocation Process

Issuers can also revoke a certificate. Smart contracts deployed on the blockchain are used to handle revocation logic, as they contain data structures to maintain the revocation status of certificates and functions to manage the process [32]. During revocation (as shown in Figure 9), using the certificate information, such as its transaction ID, the authorized issuer identifies the hash of the certificate to be revoked. Then, a revocation transaction is issued that invokes the appropriate function of the smart contract to update the revocation status [1, 10]. After that, the smart contract returns the revocation information to the DApp, through which it is sent to the issuer, and the graduate is informed about the revocation of the certificate. The DApp also updates the IPFS content of the revoked certificate to reflect the current status. This mechanism will give recruiters a means of knowing whether the certificate is revoked or not by querying the certificate status through a corresponding smart contract function.

5) The Verification Process

Graduates use the certificates received to apply for different opportunities, such as employment and admission to educational institutions. During the application, the graduates share the certificate along with its unique identifier/hash with the recruiters (e.g., prospective employers). This sharing can be done separately through a secure communication channel such as email, a messaging application, or other dedicated applications. As shown in Figure 10, after receiving a certificate and its unique identifier/hash from the graduate, the recruiter submits it to the DApp's verification interface and clicks the verification button to trigger the verification process. The DApp then invokes the smart contract function dedicated to the verification of certificates and passes a certificate identifier/hash as a parameter. The smart contract searches the blockchain ledger to find the matching hash and retrieves the associated certificate details. Then it returns this information to the DApp for further processing. Using the certificate hash/id obtained from the smart contracts, the DApp also retrieves the corresponding certificate files, issuer's profile information, and...
other metadata from the IPFS to complement verification. This IPFS information is also returned to the DApp. The DApp performs several verification checks to validate the certificate and the verification result is presented to the recruiter, indicating whether the certificate is valid, invalid, or revoked. This includes displaying the raw certificate and issuer profile information from the IPFS for recruiters to preview. Finally, the graduate is informed of the verification result through a separate communication channel. In this way, recruiters can verify the authenticity of a certificate in terms of its provenance, integrity of its content, legitimacy of its issuer, revocation status, and impersonation of its holder.

6) Handling Equivalent Statements

As stated above, according to Tanzanian regulations, educational qualifications obtained outside the country must be converted into equivalent Tanzanian qualifications. As a result, a comparable certificate is produced and issued to the graduate to be used in Tanzania, commonly known as a certificate of equivalence or equivalent statement. To comply with this requirement, the proposed model has incorporated a mechanism for issuing, revoking, and verifying equivalent statements. The process of handling these certificates is very similar to the process involved in domestically issued certificates. The only difference is that an equivalent statement is issued by the regulatory bodies (regulators) instead of the issuing institutions (issuers). To issue the equivalent statement, the regulator prepares an equivalent statement in PDF format. This document, along with other information, is uploaded to the DApp. The DApp forwards them to the IPFS, where they are hashed to generate a unique hash value. The raw files are stored in the IPFS while the hash is returned to the DApp. The DApp then sends the hash and related metadata to the blockchain’s smart contract through a transaction signed using the regulator’s private key. The smart contract must verify the source of the transaction to ensure that it comes from an authorized regulator. After confirmation, the transaction is allowed, and the hash is stored in the ledger. Finally, the transaction id/hash is returned to the DApp, presented to the regulator, and the regulator shares it with the graduate along with the equivalent statement.

To revoke an equivalent statement, the regulator begins by identifying the hash of an equivalent statement to be revoked and then triggers a revocation transaction through the DApp, which invokes the smart contract to update the revocation attribute associated with the equivalent statement’s hash. Subsequently, the revocation information is sent back to the DApp, through which it comes to the regulator, and the regulator informs the graduate of the revocation. Similarly, the DApp updates the IPFS content to reflect the current status. In the verification part, the recruiter submits the equivalent statement’s id/hash given by the graduate to the DApp and initiates the verification process. Using the identifier/hash, the DApp retrieves the equivalent statement information from the blockchain and IPFS and performs the verification checks to authenticate the document. This includes checking its revocation status and displaying the raw equivalent statement files from the IPFS for the recruiters to preview. The verification result is then presented to the recruiter, indicating whether the equivalent statement is valid, invalid, or revoked. Finally, the graduate is informed of the verification result.

D. How the Proposed Blockchain-based Solution Addresses the Identified Verification Challenges

Most of the limitations of the current verification system in Tanzania emanate from paper certificates and manual verification procedures. This model can eliminate both of them by introducing digital certificates and automating the tasks involved in the issuance and verification process. Recruiters will be able to verify certificates by a simple search on the blockchain system and easily get reliable results in real time [10, 26], eliminating burdensome, lengthy, costly, and inefficient verification processes [2, 38]. One further problem with current paper-based certificates or certificates of equivalence is that the embedded security features (e.g. holograms, signatures, etc.), which act as proof of their authenticity, are not verifiable by the recruiters because of their unfamiliarity with them and lack of specialized tools. However, with digitally signed certificates and their metadata in the blockchain, the digital fingerprint or hash value stored on the blockchain becomes their proof of authenticity [1]. This fingerprint or hash value can be easily accessed by the recruiters to verify a certificate by a simple lookup on the blockchain, making the certificates easily verifiable [11].

The traditional verification system in Tanzania has a high dependence on issuers for the verification of certificates and equivalent statements. This means that recruiters cannot reliably verify certificates and equivalent statements without consulting the issuers. This imposes a high workload on issuers because they have to mobilize resources and dedicate time to handle verification requests, while recruiters spend a lot of time and money consulting issuers. However, the disintermediation property of the blockchain removes the dependency on central authorities [86]. Therefore, the proposed blockchain-based
solution will allow recruiters to authenticate certificates and equivalent statements on their own without relying on the issuing institutions [3, 87]. This will eliminate the need to consult the issuing institutions, speed up the verification process, reduce verification expenses, and relieve workers of work overload [3, 10-11, 26, 87]. Additionally, most institutions maintain registries or computer systems to retrieve the information required to print and verify certificates or equivalent statements. These systems are based on centralized databases in combination with some physical storage and introduce a single point of failure, are susceptible to internal threats, and internal officials may tamper with records and produce fake certificates. However, due to the decentralized and distributed nature of the blockchain, where records are replicated on all nodes in a P2P blockchain network [3], it is difficult for the system to be attacked [39], fail, or lose data [11, 26, 87]. Furthermore, the immutability, tamperproof, and transparent nature of blockchain makes it impossible for internal officers to manipulate stored records and produce fake certificates [10].

Regarding the problems caused by disintegrated verification systems, the proposed blockchain-based system can solve the problem by allowing all certificates, including equivalent statements, to be issued on a common platform from which all recruiters can verify them through a single interface [38, 88]. In so doing, even hundreds of certificates and equivalent statements can be verified effortlessly, in a matter of minutes, and with less cost. In traditional verification systems, as found in the case of Tanzania, it is difficult for issuers to revoke a certificate or an equivalent statement and for recruiters to examine if it is revoked or not. In the proposed blockchain-based system, revocation can be achieved by including a revocation status in the data structure of the smart contracts deployed on the blockchain [32, 37]. Through a specific smart contract function, the status can be set to indicate that the certificate or equivalent statement is revoked [1, 10]. Therefore, anyone who tries to use a revoked certificate fraudulently will be prevented. This revocation status will be visible to recruiters during verification.

### Table 1. The Verification Challenges and Their Corresponding Blockchain-Based Solutions

<table>
<thead>
<tr>
<th>Identified Certificate Verification Challenges</th>
<th>A Blockchain-Based Solution to the Identified Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper-based certificates and manual verification procedures</td>
<td>The proposed solution introduces digital certificates and automates the verification process. The recruiters will verify a certificate or an equivalent statement by simply querying its hash on the blockchain and get results instantly.</td>
</tr>
<tr>
<td>Disintegrated certification and verification systems</td>
<td>The proposed solution introduces a single platform through which all certificates or equivalent statements from different education levels can be issued and a single interface through which the recruiters can verify them.</td>
</tr>
<tr>
<td>Dependency on the issuers for verification of certificates</td>
<td>The proposed solution leverages the disintermediation property of the blockchain that enables the recruiters to verify the authenticity of certificates and equivalent statements without relying on the issuers or regulators, removing the need to consult the issuing institutions.</td>
</tr>
<tr>
<td>Unverifiability of the current paper-based certificates</td>
<td>The proposed solution uses the hash value of a certificate or equivalent statement stored in the blockchain as their proof-of-authenticity. This hash value is easily verifiable by a simple lookup on the blockchain, hence making the certificates easily verifiable.</td>
</tr>
<tr>
<td>Challenges regarding communication between parties involved in the certificate verification process</td>
<td>With the proposed solution, communication between the parties involved in verifying certificates or equivalent statements will no longer be needed because the recruiters will verify them independently by querying their hash from the blockchain.</td>
</tr>
<tr>
<td>Vulnerability of centralized systems used to maintain certificate records</td>
<td>The proposed solution leverages the decentralization and distribution nature of the blockchain, where data is replicated to all nodes in a P2P network, making it resilient to attacks, failures, or data loss. In addition, its immutability property makes it secure from internal threats.</td>
</tr>
<tr>
<td>Certificate revocation issues</td>
<td>The proposed solution uses on-chain revocation by including a revocation status on the smart contracts deployed in the blockchain. This status can be set to indicate that the certificate is revoked and recruiters will invoke this status to tell if the certificate is revoked.</td>
</tr>
</tbody>
</table>

In general, apart from addressing the challenges of the current paper-based certificates, as shown in Table 1, the following are the main functionalities and contributions of the proposed model:

- **Addresses the certification problem at different levels of education across the country.** This can be achieved by ensuring that graduates use the same blockchain address and ID throughout their educational journey and that all acquired certificates or equivalent statements are associated with them.

- **Incorporates a hierarchical scheme of identification and authorization that includes the central authority, the regulators, and the issuers.** Using smart contracts, the central authority registers the regulators, and the regulators register their respective issuers on the blockchain. It ensures that the entities involved in transactions are known and authorized to avoid degree mills.

- **Emphasizes how to control the access permissions of different roles involved in the system.** Smart contracts will be used to incorporate information about the participating entities and the rules that define their permissions.

- **Introduces an on-chain revocation mechanism.** Smart contracts provide a data structure that can include a revocation status variable, indicating if the certificate is revoked. Recruiters can refer to that status to know that the certificate or equivalent statement is revoked.

- **Facilitates the issuance of certificates and equivalent statements on the blockchain.** The process involves hashing a certificate or equivalent statement and storing its hash on the blockchain while storing the raw files and other complementary information in the IPFS.

- **Provides a mechanism for third parties (recruiters) to verify certificates by directly querying the certificate's hash from**
the blockchain and other complementary information from IPFS.

- Provides a mechanism for issuing and verifying foreign certificates on the blockchain using their equivalent statements. The regulators, responsible for generating equivalent statements for foreign awards, will upload these statements to the system and the hash will be stored in the blockchain for verification purposes.

VI. CONCLUSION AND FUTURE WORK

This study investigated the educational certificate verification problem in Tanzania and proposed a blockchain-based conceptual model to address the identified challenges and improve the verification process of certificates and equivalent statements. The challenges facing the current verification system in Tanzania arise from manual procedures, unverifiable credentials, dependency on issuers, centralized systems, disintegrated systems, revocation, and communication-related issues. These challenges undermine the verification of certificates and equivalent statements, creating loopholes for forgeries to persist. The proposed model can solve these challenges by digitizing certificates or equivalent statements, cryptographically hashing them, and hosting their hashes on the blockchain ledgers so that third parties who wish to verify certificates or equivalent statements can simply query the blockchain. In addition, other unique contributions of this model include handling the certification problem at different levels of education, involving regulatory bodies to manage issuers of certificates, controlling access permissions of different roles involved in the system, on-chain revocation of certificates using smart contracts, and handling of the certificates acquired abroad.

In this approach, recruiters can perform verification without depending on issuers, accelerating the verification process, reducing expenses, and relieving them from the verification complexities. Unlike traditional systems, the proposed blockchain-based solution can be more reliable because it is highly secure against attacks, failures, and loss of records. It also includes an efficient mechanism to revoke a certificate or equivalent statement and inform recruiters. Most importantly, it is presumed that this model will positively contribute to the efforts to prevent certificate fraud. Although the proposed model is tailored to the Tanzanian qualifications system, it can be assimilated to other countries with similar educational structures that face the same challenges.

The primary goal of this study was not to provide a full implementation of a blockchain-based system but to unveil the challenges associated with the certificate verification process based on ground truth data from the Tanzanian setting and provide a detailed conceptual design of a blockchain-based model, underlining how it can address these problems. However, considering the need to demonstrate its applicability, the proposed conceptual model will be implemented to evaluate its viability in the issuance and verification of certificates for the detection of counterfeits. To be specific, a prototype or proof-of-concept based on the underlined scheme will be developed and deployed on a blockchain platform, such as Ethereum, that supports smart contracts. Finally, it will be evaluated and validated to demonstrate its performance from a practical point of view.

REFERENCES


J. Gustafsson, "Single case studies vs. multiple case studies: A comparative study."


M. Shaheen, S. Pradhan, and Ranajee, "Sampling in Qualitative Research, in Qualitative Techniques for Workplace Data Analysis, Hersley, PA, USA: IGI Global, 2019, pp. 25–51.


