SECURE DOCUMENT VERIFICATION SYSTEM USING BLOCKCHAIN

Oiza Salau
Computer Engineering Department
Nile University of Nigeria
Abuja, Nigeria
oiza.salau@nileuniversity.edu.ng

Steve A. Adeshina
Computer Engineering Department
Nile University of Nigeria
Abuja, Nigeria
steve.adeshina@nileuniversity.edu.ng

2021 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS) | 978-1-6654-3493-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICMEAS52683.2021.9739812

***Abstract*—Document verification is a complex domain that involves processes to authenticate original documents. Some original documents like birth certificate, university diploma, contract, certificate of occupancy, Will etc. may involve serious verification and authentication practices, because fake documents can easily be created. A skillfully generated fake document is always difficult to detect and can be treated as original. With the increase of forged documents, the integrity of both the document holder and the issuing authority is jeopardized. This research is intended to address the issue of electronic document forgery and provide an alternative secure means for storing documents. The aim of this research is to design and implement a secure document verification system using blockchain. The result of this study shows how the user documents are stored securely in the blockchain, and if any change or adjustment are made to the documents the chain becomes invalid and the user will be notified**

Keywords— blockchain, decentralization, hashing, interplanetary file system, digital certificate

I. INTRODUCTION

In today's world, digital documents are now becoming a large part of every section whether private or public resulting from the change in modern technology. This does not only allow the spreading of information but also saves the documents in digital form and helps promotes a paperless environment [1]. Digital documents are easy and convenient to use but, proving their authenticity is often difficult. Since there is no common protocol to verify and validate digitized documents, the process of analyzing the documents in government and private institutions is difficult [1].

In a country like Nigeria, the document verification process is still based on printed documents. Original documents such as, insurance documents, wills, academic certificate, contracts, birth certificate, marriage certificate, case files etc. these documents might be subjected to a process of verification by an authority using signature, and this process is known to be time consuming and cumbersome. Later if there is a need for another verified copy of the same documents, the whole process will be repeated or a copy of the document will be verified by a third party. In addition, there could be loss of some documents due to mismanagement, a greater challenge is experienced with the advancement in technologies, which makes it possible to reproduce falsified and fake document, as they are now available to unauthorized alterations.

A number of forgery cases have been recorded and reported in the past. In Nigeria, a former acting director general of the federal institute of industrial research Oshodi. Claimed to have gotten a PhD from a university in Benin Republic which turned out to be a fake certificate after being investigated by the Independent Corrupt Practices and other related Offences (ICPC) [2], Similarly a senior lecturer at Nnamdi Azikiwe University Awka, was found with a fake master's degree and a plagiarized PhD [2]. Also, in South Africa some individuals presented a fraudulent asylum document at their work place which they claimed to have received from the officials of the state agency [3]. In addition, a woman in India who was requesting for an international passport submitted a fake certificate of birth. The police said it is a common occurrence when individuals apply [3]

The counterfeiting of these documents allow individuals to get credit for what they do not deserve. Cases such as this shows the severity of document fraud. It is important to get a solution that will solve this problem as required, by introducing different methods to secure and safe guard the original documents from being forged.

The downside in securing documents has attracted the attention of many researchers and the research in this field keeps improving, with the sole aim to stop unauthorized alteration and compromising the integrity of this document. There are different methods that are used to add extra information to the authentic documents to ensure that a copy will not be altered in anyway. Some of these techniques are barcode, hashing, document signatures, watermarks etc. [4] However, the information stored using these above listed methods have their limitations, for example using a barcode to secure a document has space limitation because all the content of a document cannot be included in the barcode. In hashing, instead of storing the original content, the hash of the document will be stored.

Regardless of this, hash values are reliable solution when using blockchain based method, this is now applied to document verification [4]. The blockchain has made it possible to implement solutions that verify the integrity of documents issued. Blockchain is a distributed, synchronized and replicated public ledger. In this research the use of blockchain technology along with interplanetary file system (IPFS) will be incorporated to present a solution that will focus on the ability to validate and verify the integrity and also store the content of an electronic document in a system.

II. LITERATURE REVIEW

This section discusses the types of method that can be used in document verification, it also discussed blockchain technology and comparison between blockchain and digital signatures, the importance of IPFS in blockchain and some related works in this field.

With the advancement, of technology over the years it has made it easy to produce digital documents that are easy to access, store and retrieve. It has also encouraged the change to a more paperless environment. Although, the issue of verifying this documents remains. Several researchers have been trying various methods and schemes to overcome this issue. Some methods are used to add extra information to the original documents to make sure that a copy will not be altered in any possible way. Some of these methods are watermarks, barcode, document signatures, hashing etc. In this research I will be discussing two popular methods that could be used for document verification.

i. Digital signature with Public Key Infrastructure PKI
ii. Blockchain Technology

A. Digital Signature with Public Key Infrastructure

Public Key Infrastructure (PKI) is a technology for authenticating devices and users in a digital world. The idea is to have one or more trusted parties to digitally sign a document certifying that a particular cryptographic key belongs to particular device or user. The key can be used as an identity for the user in digital networks [5]. Trusted Third Parties (TTP), are known as a secured middle layer on cloud service transactions they allow secure, trustful, interaction between two parties. Certificate Authority (CA), are a type of trusted third parties (TTP) that delivers validation authority for a PKI [6]. A certificate authority is a company that acts to validate the entities such as email addresses, websites, companies or individual users and bind them to cryptographic key through the issuance of digital certificates. A digital certificate provides [5];

authenticity**,** by serving as a credential to validate the identity of the entity is issued to.

Encryption*,* for secure communication over insecure networks like the internet.

Integrity, of documents signed with the certificate so they cannot be altered by a third party in transit

B. Blockchain Technology

Blockchain technology has been in existence since the year 2009, since it has been existing for eleven years, it has succeeded in resisting attempts to take it down , hack into it or corrupt it [7]. Researchers are always searching for places where blockchain technology can be implemented. Blockchain can be defined as a publicly accessible ledger in a distributed network that records digital transactions. The main strength of blockchain is the ability to provide a single source of truth through its characteristics which are immutability, decentralization, security and transparency [8].

The blockchain is a type of distributed public key infrastructure, there are some differences between the application and purpose of digital signature in a centralized conventional public key infrastructure and the application and purpose of digital signatures in the blockchain. In the blockchain, the hash value authenticates both block data and transaction data. While in public key infrastructure, the authentication of the digital certificate is done using the hash value. In addition, the application that generates the blockchain hash value can be stored privately and separately. [9]

The blockchain adaption is more advantageous to documents with a permanent retention schedule. This is because the hash value, which is a main feature of the blockchain, stands as validation metadata that would not require specific software for future verification. In addition, the blockchain records does not require a centralized network such as the certificate authority to notarize or verify the hashes.

C. Interplanetary File System

Inter Planetary File System, is a peer-to-peer distributed file system, that helps to connect all computing devices with the same system of files. The IPFS distributed network uses BitTorrent and Git technologies [10]. It is a protocol that allows users to share files and information efficiently across the network. It can be used to transfer any file from text to videos, it is ideal for sharing large files that may require or consume large bandwidth. The files on the IPFS has a unique hash values, which serve as the fingerprint for the file. IPFS can be used with blockchain, by allowing one access large amount of data and place the immutable permanent link into the blockchain transaction, this help to timestamp and secure the content without having to put the data itself in the blockchain [11].

D. Related Works

The author in [1] designed a paper based document authentication system using digital signature and QR code. The message and digital signature are compressed into the QR code, then a message along with the QR code is printed on a paper. On the receiving end after uncompressing the data, the digital signature will be verified by comparing the hash value of the obtained message and the hash from the decrypting the signature. If both hash signature is correct, then it is valid. To check the printed message, they used optical character recognition OCR. The hash value of the message gotten from the OCR is computed and compared with the hash obtained from the message and QR code. If the hash is the same, then the printed message is authentic. However, if they are different it cannot be concluded that the message is altered except after human review.

The author in [2] presented a solution for document verification using text extractor, digital signature and correlation score. Their system was designed based on three modules. Data extraction module, this module creates a machine readable format using JSON, the owner or a third party signs the document. The data validation module, a third party who can validate the integrity of the first signature also signs the document saying the first signature is valid. Score calculating module, this module is used by a company or an individual who wants to verify the document, based on the amount of signatures on the documents he verifies it.

Another author [12] worked on blockchain based framework for educational certificate verification. The paper focused on how to improve the confidentiality, ownership and authentication on an education certificate using hyperledger fabric framework. Their aim was only on academic certificates.

The author in [13] worked on enhancement of QR code capacity by encrypted lossless compression technology for verification of secure e-document. They worked on a novel method to improve the storage capacity of a quick response code. The aim of their design was to store more encrypted data or message in a QR code compared to the normal storage capacity that a QR code works on without loss of data.

The author in [14] designed a system for document verification with an effective expansion to enhance the MD5 hash functions. Their focus was on how to improve the security of the MD5 algorithm against attacks like dictionary attack, rainbow table attack and brute force attack.

The author in [15] they designed a document verification system using. They were more concerned about security so they proposed a two level QR code with elliptic curve digital signature algorithm ECDSA. The QR code stores the data that was encrypted, a strong encryption algorithm used along with strong key exchange algorithm which contributes to the higher security to the data stored in the codes.

The author in [16] designed a system for digital document verification using blockchain. The system was designed to review the certificate and personal ID from the university and ID card verification site. If the students ID cards and certificate are original, it saves the serial number of the certificate and the ID on the blockchain. Then the system generates a QR code that is kept on the e-certificate along with a serial number will be sent to the user. A company can now verify the certificate using the serial number on the blockchain or by scanning the QR code.

The author in [17] designed a certificate verification system using blockchain. The aim of their research was to eliminate educational certificate fraud with the use of blockchain technology. Their proposed system worked by registering a university in the blockchain using wallet. Once the university is added, the university will create the certificate using data fields and store the hash on the blockchain, the resulting transaction ID will be sent to the student. With the use of that ID, anyone would be able to verify the document on the system blockchain.

The reviewed literatures revealed that several studies, research, practical works have been done to enhance document verification system. With the above review some literature gap was revealed, most of the reviewed literature design made use of QR codes and digital signature for verifying the documents. The few researchers that used blockchain technology for document verification did not incorporate IPFS for storing the document before hashing and storing the immutable permanent link in the blockchain. With the use of IPFS the large document would be stored without having to worry about space.

In this project the use of blockchain technology along with interplanetary file system (IPFS) will be incorporated to present a solution that focuses on the ability to verify the authenticity, integrity, nonrepudiation of an electronic document and also store and share the content of an electronic document in a secure system.

III. PROPOSED DESIGN FOR THE DOCUMENT VERIFICATION SYSTEM

In this research, a secure document verification system using blockchain was developed based on relevant technology. The system is a web based application that was programmed using java. The system uses a self-developed blockchain written in java language.

To create a secure document verification using blockchain. All organizations / users will have an address from which it can send transactions. Users can only be added by the system admin, once added the user can have access to the system. Each uploaded documents will be stored in the interplanetary file system (IPFS) which will return a unique hash of the documents. The hash serves as a unique identity for each documents and will be encrypted along other information like name and date created using SHA-256 encryption algorithm. The result of encryption is stored on the blockchain and a resulting transaction ID will be issued to the user. With the ID anyone can have access to verify the document and view the document details. With this it is almost impossible to create a counterfeit of a document using the same data.

The use case diagram in figure1 below shows the interaction between the user and the system. A user can create account, upload document, and view document. the document the user uploads will be stored in the IPFS before stored in the blockchain. After verification of the document by the system which was uploaded, the system will send an ID to the user. With this ID the user can check the verification of the document anytime and anywhere in the world.

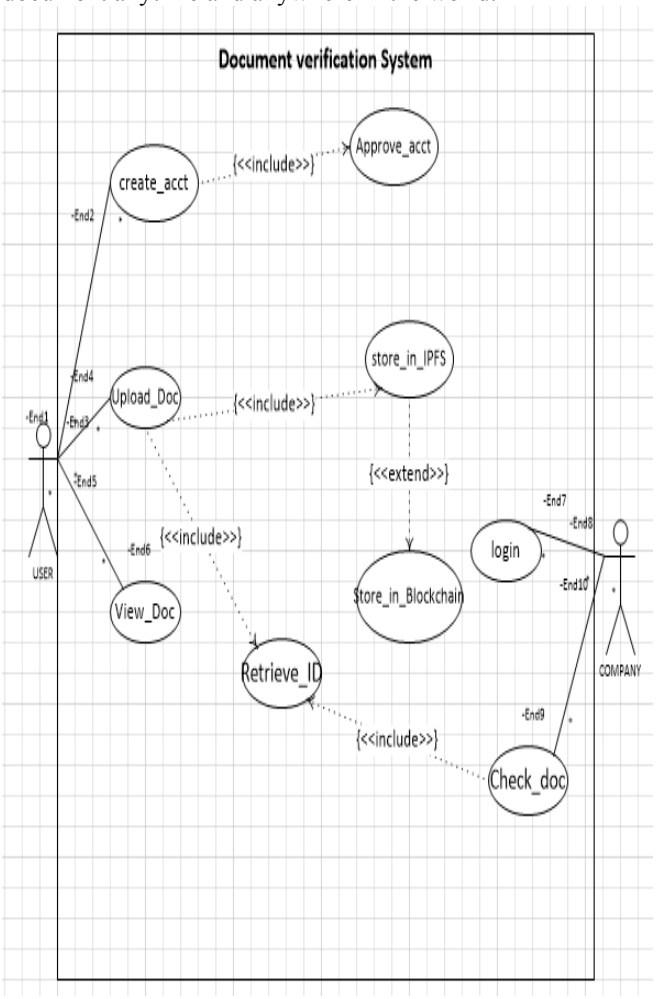

Fig1. Use Case Diagram for the Document Verification System

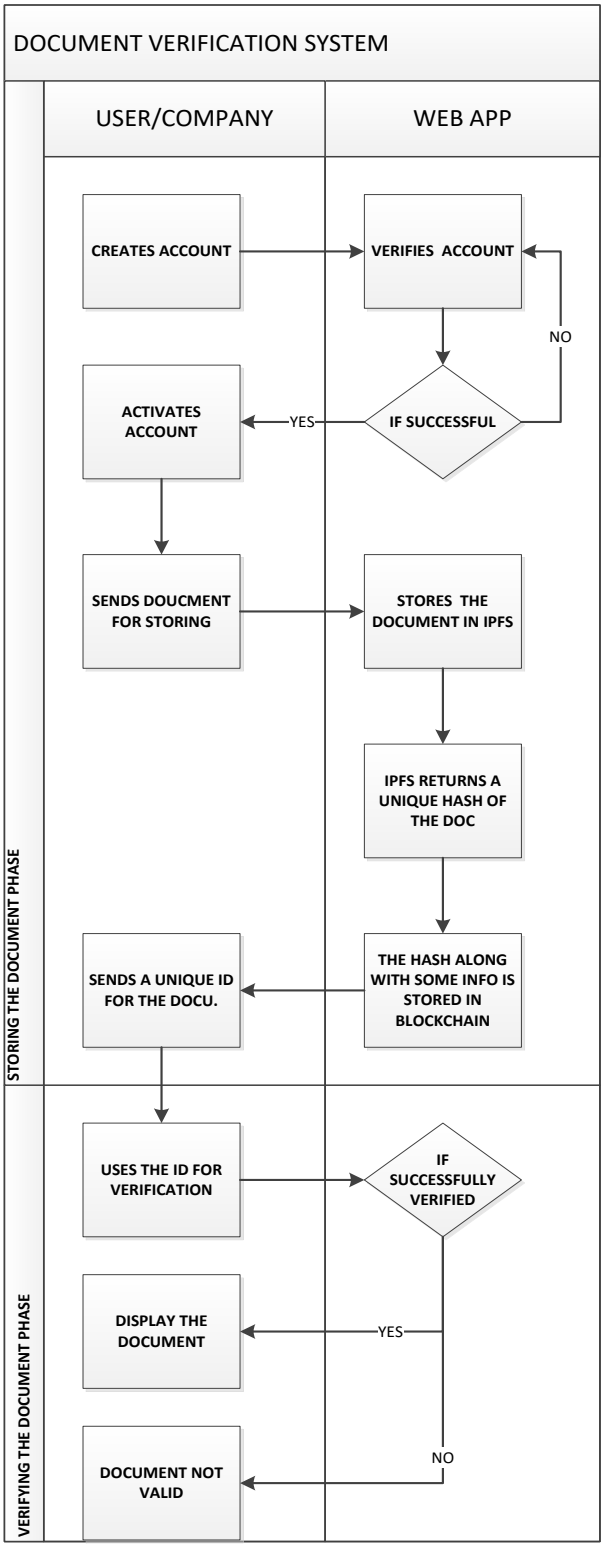

Fig 2. Flow chart diagram of the system

Figure 2 shows the flow chart diagram of the system. It shows systematically how each processes are executed. Starting from the creation of account, which is done at the user end to the verification of the document, which is done at the back end of the system.

A. Pseudocode

```
Algorithm for Document Verification Using Blockchain
Methods:
Register User
Login
Save Document
Get Documents
Download Document
Validate Chain

RegisterUser Method:
   Get User Details
   Set User Authority
   user_exist = Check Username does not exist
   if user_exist:
      return error of username exist
   else:
      Encode user password
      Save User on portal_user table

Login Method:
   Check portal_user table for Username
   login_success = Check password matches encoded password for user && user is active
   if(login_success):
      update user last login
      generate token for user
      return user token
   else:
      return error of invalid login

HashDocumentInfo Method:
   convert document object to json string
   hash string with SHA-256 algorithm
   return newly generated hash string

SaveDocument Method:
   Validate Token and Get User Detail
   set uploaded by and upload date
   store file on ipfs and get ipfs hash
   last_doc = order by id descending get first item on ordered list for table document_info
   set last_document hash on new document
   set document hash with HashDocumentInfo()
   store document info on document_info table

GetDocuments Method:
   Validate Token and Get user Detail
   search document_info for document with user
   return list of document

GetChain Method:
   Select Document Hash column from document_info ordered by id asc
```

```
ShowChainValidation Method:
   select document info from document_info ordered by
id ascending
   Create response as string array size of documents list
and pre fill with 'not valid'
   for i = 1 to array length:
      current_info = info on position i
      previous_info = info on position i-1
      if current_info hash not equal
HashDocumentInfo(current_info):
         return response
      if current_info previous_hash not equal
HashDocumentInfo(previous_info):
         return response
      set response(i) to 'is valid'
   return response
```

IV. RESULT AND DISCUSSION

In this section, the implementation of the system is discussed and the technical information about the system is presented. IntelliJ is the IDE that was used in writing the software in java. Springboot application was used as the framework for developing this web application. The application has an IPFS for storage, a database for storage, a service layer which is written in java. A frontend written in HTML, CSS, and JavaScript.

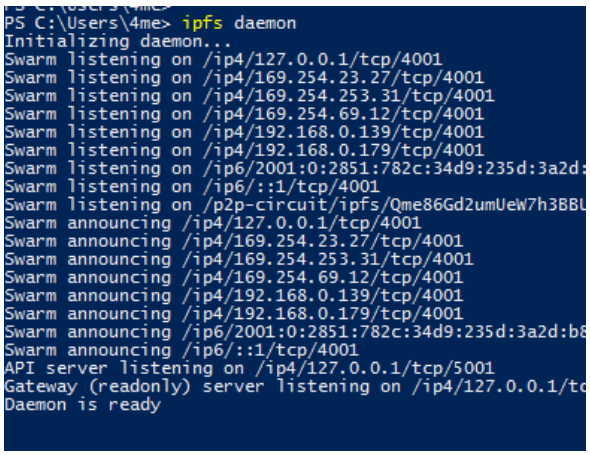

Fig 3. Starting the IPFS

Figure 3 above, shows how the IPFS starts it scans through all necessary dependencies before the application starts.

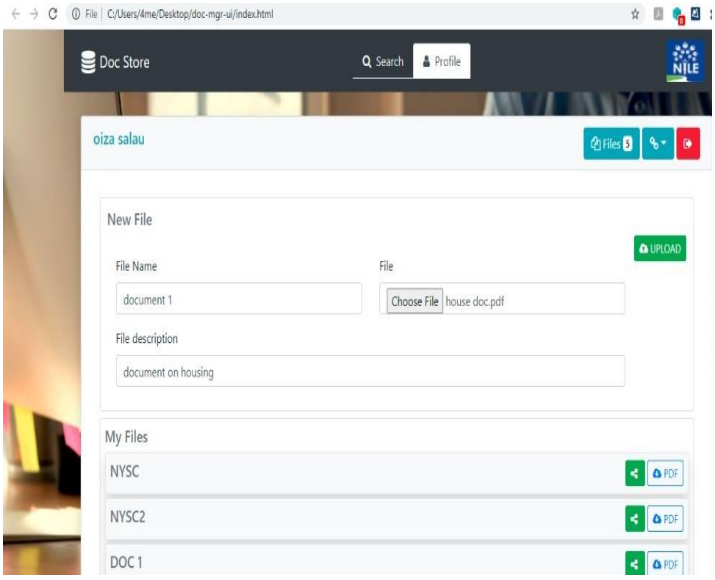

Fig 4. Saving a file on the system

The image above shows how a user saves a document on the system. The user fills in the required space and uploads his document. when the user clicks on the upload button the document will be saved and the document ID which is the hash will be sent to the user as shown in figure 5.

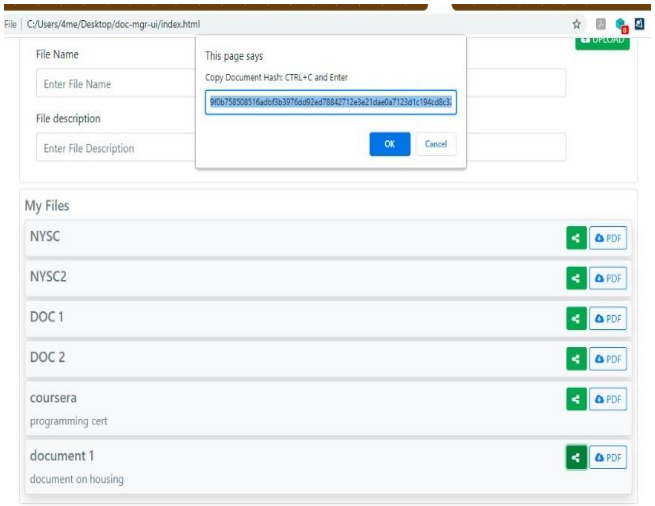

Fig 5. List of stored document on the system

Figure 5 shows the list of stored document on the system. If a user wishes to share a file. The user clicks on the "share" button and the hash value for that file is displayed as shown above. It is the hash value that the user shares. With the hash the document can be displayed and verified from anywhere in the world.

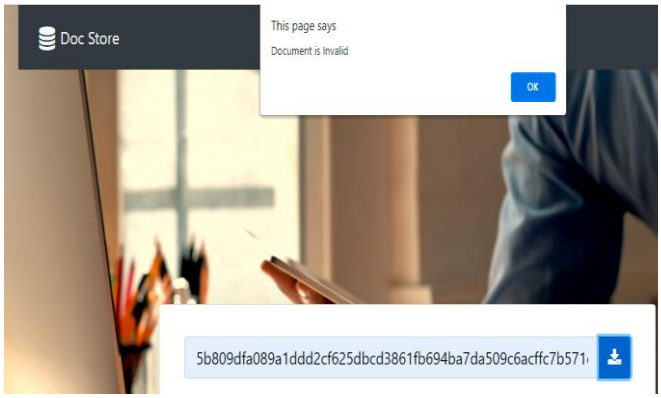

Fig 6: verifying an altered document

The image above shows, what happens if an altered document is being verified. An error message "document is invalid" will be displayed on the screen. When the system checks for the validity of a document it checks the content of the document with the existing content of the document stored in the database.

Comparing this system with other document verification system that researchers have worked on. This system uses IPFS to store the documents, while the hash is stored on the blockchain. This allows the storage of the system to be large and also increase the efficiency.

V. CONCLUSION

The challenges surrounding document forgery and also to design a certification system is the key issue this research has set out to resolve. Although blockchain has been around for a while, research into its application are just beginning to find expression within the Nigerian context. This system was developed using java programming language and MySQL database. Upon launching the application, the user creates an account with the necessary information. After which a user can now upload a document to be stored, the document is first stored in the IPFS which hashes the document and then stores the hash in the blockchain. Also a user can verify a document using the document ID.

In brief, the result of this study shows how the user document is stored securely in the blockchain, and if any change or adjustment are made to the documents the chain becomes invalid and the user will be notified.

VI. REFERENCES

[1] l. Chanakal, D. Sachitra, H. Chandru, W. Chamin, G. Dias and S. Fernado, "IDStack The common protocol for document verification built on digital signatures," *IEEE Xplore,* pp. 8-12, 2017.

[2] O. Jide, "Rising Cases of Certificate Forgery in Nigeria," 4 March 2020. [Online]. Available: https://punchng.com/rising-cases-of-certificate-forgery-in-nigeria/. [Accessed 3 May 2020].

[3] M. Sthambile, D. Nelisiwa and G. Barbour, "Proposing a Blockchain-based Solution to Verify the Integrity of Hardcopy Documents," *council for scientific and industrial research (CSIR),* vol. VII, pp. 19-24, 2018.

[4] T. Maurizio, A. Franco and D. Andrea, "A blockchain based PKI Validation System based on Rare Events Management," *future internet,* vol. 5, pp. 3-6, 2020.

[5] DocuSign, "Understanding Digital Signature," New York, 2020.

[6] S. Thompson, "the preservation of digital signatures on blockchain," University of British Columbi ischool, Columbi, 2020.

[7] J. Rayan and D. Daniel, "VisualDiff: Dcument Image Verification and Change Detection," *12TH International Conference on Document Analysis and Recognition,* pp. 21-24, 2013.

[8] H. Ysn and W. Caifen, "A Novel Blockchain Based authentication Key Exchange Protocol and Its application," *IEEEXplore,* pp. 20-23, July 2018.

[9] O. Saleh, O. Ghazali and M. E. Rana, "Blockchain Based Framework for Educational Certificate Verificatio," *Advance Science Institiue,* pp. 19-23, march 2020.

[10] N. Emmanuel and P. Reza, "BlockIPFS- Blockchain-enabled Interplanetary File System Forensic and Trusted Data Traceability," *IEEE International Conference on Blockchain,* vol. 8, no. 12, pp. 17-19, 2019.

[11] N. Nizamuddin and S. Hasan, "IPFS-Blockchain Based Authenticity of Online Publication," june 2018. [Online]. Available: https://www.researchgate.net/publication/325899234_IPFS-Blockchain-Based_Authenticity_of_Online_Publications.

[12] K. Ntin and S. Mengade, "Certificate verififcation system using blockchain," international journal for research in applied & engineering technology (IJRASET), APRIL 2019. [Online]. Available: www.researchgate.net.

[13] W. Maykin and K. Pramote, "Paper based document authentication using digital signature and QR code," *4TH International Conference on Computer Engineering and Technology (ICCET),* vol. 4, pp. 12-17, 2012.

[14] B. Expert, "Blockchain Timestamp and Document Verification," 2019.

[15] w. Maykin and P. Kuachareon, "Paper based document authentication using digital signature and QR code," *international conference on computer engineering and technology (ICCET),* vol. 5, pp. 16-24, 2012.

[16] A. Momot, "How Blockchain addresses PKI shortcoming," 2017.

[17] A. Ramon and C. Bellver, "Blockchain in Education Introduction and Critical Review of the state of Art," *Revista Electronica De Technologia Educativa EDUTEC,* vol. V, pp. 41-44, November 2017.

[18] A. Wright and F. Daniel, "Decentralized Blockchain Technology and Rise of Lex Cryptographia," *SSRN Electronic Journal,* pp. 17-19, March 2015.