Major Project

Decentralized Storage Of Documents Using NFTs And Block chain Technology

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF TECHNOLOGY
in
INFORMATION TECHNOLOGY

by

Neeraj Kumawat
222IT024

under the guidance of

Dr. Dinesh Naik

DEPARTMENT OF INFORMATION TECHNOLOGY
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE - 575025

November, 2023

DECLARATION

I hereby *declare* that the Report of the P.G. Project Work entitled Decentralized Storage Of Documents Using NFTs And Blockchain Technology, which is being submitted to National Institute of Technology Karnataka, Surathkal, in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Information Technology in the Department of Information Technology, is a *bonafide report of the work carried out by me.* The material contained in this Report has not been submitted at any University or Institution for the award of any degree.

Place: NITK, Surathkal
Date:24/11/2023

Neeraj Kumawat
(222IT024)

CERTIFICATE

(for internal)

This is to *certify* that the P.G. Project Work Report entitled Decentralized Storage Of Documents Using NFTs And Block chain Technology, submitted by Neeraj Kumawat (Register number: 2220203) as the record of the work carried out by him, is *accepted as the P.G. Project Work report submission* in partial fulfillment of the requirement for the award of degree of Master of Technology in Information Technology in the Department of Information Technology.

Dr. Dinesh Naik
Project Guide
Dept. of Information Technology
NITK Surathkal, Mangalore

Chairperson-DPGC and Head
Dept. of Information Technology
NITK Surathkal, Mangalore

ACKNOWLEDGEMENT

I have taken efforts in this project, however, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend my sincere thanks to all of them. I am highly indebted to Dr. Dinesh Naik for his guidance and constant supervision as well as for providing necessary information regarding the project also for her support in completing the project. I would like to express my gratitude towards all members of National Institute of Technology Karnataka Surathkal, Information Technology Department for their kind cooperation and encouragement which helped me in completion of this project

Neeraj Kumawat

(Reg. No.: 2220203)

ABSTRACT

In today's world,documents is an integral art of the societal life. Where the user records are vulnerable to mutability, hacking and tampering due to the centralized storage of the assets. Addressing these problems, this paper aims to solve them by the application of a decentralized model using the Blockchain Technology. The user's assets are cryptographically hashed and represented using the Non-Fungible Tokens(NFTs) to uniquely identify each asset, which are stored on the Ethereum Blockchain, Due to the limitations of cost and data storage on the blockchain, Interplanetary File System (IPFS) is implemented, where the original data is stored and the unique Id returned by the IPFS i.e., Content Identifier(CID) is used by the users to access the Asset.

Keywords — IPFS,Blockchain,HTML,CSS,JS,ReactJS,Decentralization,Solidity,Firebase

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

All individuals possess a variety of documents ranging from identity proofs, educational certificates, official documents, bonds, declaration certificates, confidential documents. Companies, Educational institutions, government bodies, medical officials need to store, issue, share, dispatch documents for public benefits. Storing them digitally or issuing a digital copy has become a common practice all around the world as a sustainable way. With this gaining more popularity, it becomes crucial to organise and safeguard them for easy and timely retrieval. Most widely accepted way of storing the documents is storing them on centralized systems such as Google Drive. According to [1], Google Drive has 1 billion users worldwide. However, centralized storage of data is often susceptible to data corruption, tampering and failures and is not the most trusted way for data security. Besides Google Drive, another widely used third-party application is Whatsapp. Several researches [2], have proved that the end-to-end encryption used by this app can be hacked easily, thus calling for an alternative.

Decentralizing the data using Blockchain Technology is being looked at as a potential solution to all these challenges. Blockchain Technology has found applications in almost all sectors such as real estate, healthcare, trade and finances, logistics etc. and has been growing ever since. Blockchain forms an essential part of one of the most trending topics crypto-currency since the best known currency Bitcoin makes use of this technology. Apart from bitcoin, the second-largest crypto-currency Ethereum is an open source Blockchain. Multiple advantages provided by Blockchain such as improved transparency through use of distributed ledger, high efficiency, reduction of costs, easy trace-ability and immutability of data. All of these features and more make it a perfect solution for data management and storage. Since it facilitates storing data on more than one or multiple servers, this technology is highly reliable and immune to failures. In broad terms, technology collects information in groups, where each group is known as a block and each of these blocks contain a set of transactions which are encrypted and shared in a deeply secure way among multiple servers thus also ensuring confidentiality and integrity.

1.1 Challenges

Integrating a face recognition system for extracting images from documents and matching them with machine learning algorithms like K-Nearest Neighbors (KNN) presents the challenge of achieving high accuracy, influenced by the quality and diversity of training data, algorithm selection, and minimizing false positives. Simultaneously, creating a user-friendly QR code system for document uploads on IPFS like to UPI payments demands a seamless user experience, secure integration with IPFS, privacy considerations, scalability planning, and robust error handling to ensure a smooth and reliable platform

1.2 Motivation

Your motivation to build this platform is commendable. Addressing document forgery and ensuring the security and integrity of important documents is a significant societal and organizational challenge. By creating a platform that leverages technologies like blockchain, face recognition, and IPFS, you are contributing to enhancing document verification and security, which can have far-reaching benefits in areas such as identity verification, education, and legal processes. This endeavor not only has the potential to simplify and strengthen document management but also promotes trust and transparency in various sectors where document authenticity is critical. It's an important step towards ensuring the reliability and tamper-resistance of vital records, ultimately making processes more efficient and trustworthy.

Chapter 2

Literature Review

2.1 Background

In today's digital era, the significance of documents, encompassing a multitude of types such as educational certificates and critical records, cannot be overstated. However, traditional centralized databases have proven susceptible to unauthorized alterations, raising concerns about document security and authenticity. To address these challenges, our vision is to introduce innovative technology solutions. We are combining the power of blockchain and IPFS (InterPlanetary File System) to revolutionize document storage and verification. By utilizing blockchain for its immutable ledger capabilities and IPFS for efficient and decentralized storage, we aim to establish a robust ecosystem where each document is securely stored and uniquely identified through Content Identifiers (CIDs). This breakthrough approach not only ensures the integrity of documents but also promotes trust and accessibility. Our motivation stems from the need to combat document forgery and create a tamper-resistant system that can enhance transparency, reliability, and efficiency across various domains. Through this initiative, we aspire to harness technology for the greater good, safeguarding the authenticity of essential documents in our increasingly digital world.

2.2 Related Work

Ethereum provides a blockchain with a built-in fully fledged Turing-complete programming language that can be used to create "contracts" that can be used to encode arbitrary state transition functions, allowing users to create systems with starkly different functionalities simply by writing up the logic in a few lines of code [3].

Blockchain has gained popularity in recent years, finding applications in numerous industries including but not limited to music, finance, construction, banking, pharmaceuticals and Internet of Things (IoT).

Das et al. described a distributed construction document management system using Blockchain and distributed content-addressable storage technologies. A blockchain ledger data model for tracking workflows and document version management is proposed. Their framework also deploys peer-to-peer content-addressable storage for preventing single points of failure and data integrity of documents through data partitioning, data replication, and cryptography.

Shrivastava et al. proposed a decentralized private Blockchain system using a private IPFS database server for storing documents over Blockchain considering the Ethereum Blockchain ecosystem. It envisaged a government owned private Blockchain where universities, students, and third parties will work as stakeholders. However it does not provide a practical implementation of the proposed system to test performance.

Nizamuddin et al. proposed a blockchain-based framework for document sharing and version control to facilitate multi-user collaboration and track changes in a decentralized manner. They addressed the security issue of document storage and proposed a mechanism for one-to-one document sharing. However, they did not describe a mechanism to share documents with more than one person at a time, thus facilitating group sharing and management of documents.

Haveri et al. proposed the EduBlock information system, a multi-node private blockchain network using the Ethereum framework and private Interplanetary File System (IPFS) to store documents. They evaluated the performance of the system by analysing the impact of various parameters such as varying difficulty level, load, and consensus algorithms.

Taufiq et al. proposed a crypto-governance graduate document storage system to solve the critical problem of fraud avoidance diplomas, transcripts and diploma supplement model in Muhammadiyah Tangerang University. Indonesia.

Ramachandran et al. proposed a blockchain based framework using IPFS and traditional encryption methods to create a scientific data provenance management framework for automatic verification of data records. The proposed platform would be used to facilitate trustworthy data provenance collection, verification and management.

Bocek et al. presented a startup modum.io that uses IoT (Internet of Things) sensor devices leveraging blockchain technology to assert data immutability and public accessibility of temperature records, while reducing operational costs in the pharmaceutical supply-chain.

Rajalakshmi et al. proposed a blockchain-based framework using IPFS and other traditional encryption methods to create a secure, tamper proof model of academic research record keeping with access control methods. Their system utilizes Ethereum smart contracts to store the provenance metadata information retrieved from the IPFS file system to the blockchain network, to create tamper-proof records for further auditing purposes.

Destefanis et al. advocated the need for a new branch of study - Blockchain Software Engineering which addresses issues posed by smart contract programming and other blockchain applications. They discussed how established best practices could be adapted to blockchain to prevent harmful software misbehavior.

This paper describes a real-world implementation of a Blockchain based education document storage system using Interplanetary File System (IPFS) based on proof of stake consensus and illustrates its benefits in terms of speed and reduction of verification steps

2.3 Outcome of Literature Review

1. IPFS for Document Storage: The research papers consistently emphasize the use of the InterPlanetary File System (IPFS) for storing documents. IPFS provides a decentralized and distributed storage solution that ensures data availability and integrity.

2. Verification Techniques: To enhance document security and authenticity, various verification techniques are employed: - **Cryptography:** Cryptographic methods are utilized to protect the confidentiality and integrity of documents, enhancing their overall security. - **Digital Signatures:** Digital signature mechanisms are applied to verify the authenticity and origin of documents, serving as a trust-building component in document verification processes. - **Machine Learning Algorithms:** Machine learning algorithms, particularly for face recognition, play a pivotal role in the verification process. These algorithms contribute to the validation of individuals'

identities against document data.

3. QR Code Generation: Researchers consistently implement QR code generation schemes for each document. This approach facilitates easy document access and retrieval, allowing anyone to scan the QR code to obtain the associated document securely.

Collectively, these findings underscore the significance of leveraging decentralized storage, advanced cryptographic measures, digital signatures, machine learning, and QR code technologies to enhance document security, integrity, and accessibility in diverse applications.

Table 2.3.1: Summary of Literature Survey

Authors/year	**Methodology**	**Observations**
Divya Shree/2021	This paper represents that implementation of verification of documents using live face recognition with KNN nearest neighbours.	Verification process becomes more secure.
Bagas Fadillah Islamay, 2022	This paper represents how to implement a document management system using Ethereum blockchain in Rinkeby test network and InterPlanetary File System or IPFS	They are basically using IPFS tech to store doc which is very safe because of documents are stored in crytpo hashed manner and Content Identifier is generated.
Ajay Kumar Shrivastava, 2021	This paper represents that we are using private IPFS database server for storing our documents over Blockchain. We are considering Ethereum Blockchain ecosystem in our case	Using a private IPFS database server alongside the Ethereum Blockchain ecosystem offers a secure and decentralized approach to document storage.
Iftekher Toufique Imam, 2021	.This paper represents that we are using private IPFS database server for storing our documents over Blockchain. We are considering Ethereum Blockchain ecosystem in our case	Private IPFS is more secure than simply storing onto IPFS
Nayana N. Kumar, 2022	This paper represents that how can we store documents in IPFS	It have used NFT concept for storing doc

2.4 Problem Statement

The current storage system for important documents, which relies on centralized repositories, is plagued by significant vulnerabilities, including fraud and document discrepancies.. To address these critical challenges, there is a need for a robust and secure document storage solution. To solve this issue, we propose the implementation of blockchain technology coupled with the use of cryptographic hashing within the InterPlanetary File System (IPFS) for document storage and retrieval.

2.5 Objectives

1. Building a User Portal for Storing Documents using IPFS.
2. Document Verification through Face Recognition.
3. Generating QR Codes for Documents.
4. User-Specific QR Code for Uploading to IPFS.

Chapter 3

Methodology

3.1 Proposed Workflow

3.1.1 Project Setup

- Set up a development environment with React.js and Bootstrap for frontend development.
- Create a Firebase project and configure Firebase authentication.

3.1.2 User Authentication

- Develop a Login Page:
  - Create a user-friendly login page using React components and Bootstrap styles.
  - Implement Firebase Authentication to enable users to log in with their email and password.
- Implement a Registration Page:
  - Design a registration page where new users can create accounts.
  - Integrate Firebase Authentication to handle user registration.
  - Ensure validation and security measures to prevent unauthorized access.

3.1.3 User Dashboard

- Create a user dashboard where authenticated users can manage their documents and perform actions.

3.1.4 Document Upload

- Develop a document upload feature:
  - Allow users to securely upload documents to the system.
  - Implement validation to ensure document file types and sizes meet requirements.

3.1.5 Face Recognition

- Integrate Face Recognition:
  - Use a face recognition library or API to compare the extracted image from the document with the user's live face.
  - Allow document upload only if the face recognition check is successful.
  - Ensure the face recognition process is secure and privacy-compliant.

3.1.6 QR Code Generation

- Generate QR Codes for Documents:
  - Develop a feature to generate QR codes for each uploaded document.
  - Embed relevant document metadata or a link within the QR codes.
  - Store the generated QR codes and associate them with their respective documents in your database.

3.1.7 User-Specific QR Codes

- Generate User-Specific QR Codes:
  - Create unique QR codes for individual users.
  - Link each user's QR code to their specific document upload location.
  - Ensure that scanning a user's QR code prompts the uploader to add a document to the associated location.

3.1.8 Testing and Quality Assurance

- Thoroughly test each feature and component of the system:
  - Perform unit testing and integration testing to ensure functionality.
  - Conduct usability testing to ensure a user-friendly experience.
  - Test for security vulnerabilities and implement fixes as needed.

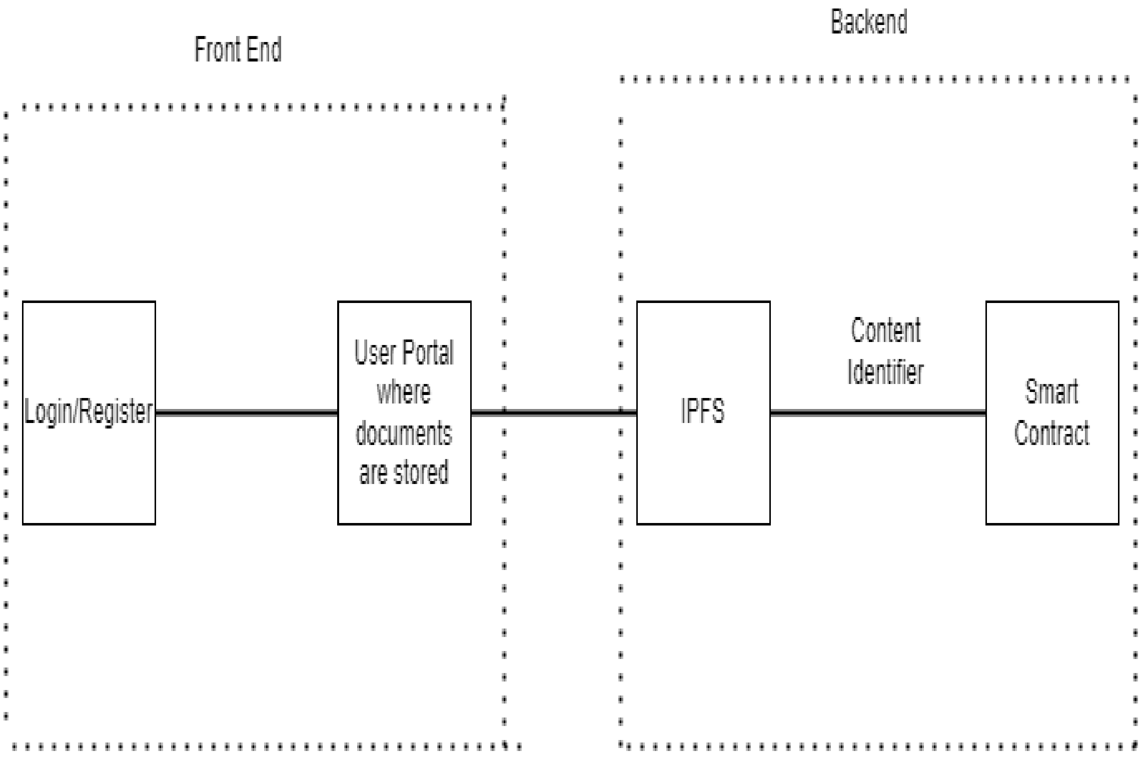

Figure 3.1.1: Proposed Workflow

Chapter 4

Work Done

1. **User Authentication System:** I have implemented a user authentication system using Firebase to ensure secure user access to the platform.

2. **User Interface:** I have developed the user interface using ReactJS and Bootstrap, providing an intuitive and user-friendly experience for storing and managing documents.

3. **Document Retrieval:** Users can download documents using a Content Identifier (CID), which enhances document retrieval efficiency.

4. **QR Code Integration:** Each document is associated with a QR code, allowing users to scan it for easy document retrieval and download.

5. **Solidity Contract:** I have created a Solidity contract for uploading documents, ensuring a secure and transparent process for adding documents to the platform.