INSTITUTIONAL REPOSITORY SYSTEM

A Major Project

Submitted in partial fulfillment of the requirement for the award of the degree of

Bachelor of Technology

In
COMPUTER SCIENCE AND ENGINEERING

By

Ayush Dobhal (Registration No.: 11019210066)

Harsh Gupta (Registration No.: 11019210081)

Sumeet Jain (Registration No.: 11019210105)

Under Supervision of

Dr. Puneet Goswami
(H.O.D. Dept. of Computer Science and Engineering)

Ms. Lakshita Aggarwal
(Asst. Professor Dept. of Computer Science)

SRM
UNIVERSITY
DELHI-NCR, SONEPAT

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
FACULTY OF ENGINEERING SRM UNIVERSITY DELHI-NCR

CERTIFICATE

This is to certify that the project entitled "**INSTITUTIONAL REPOSITORY SYSTEM"** submitted by Ayush Dobhal, Reg. No. 11019210066, Harsh Gupta, Reg. No. 11019210081 and Sumeet Jain, Reg. No. 110192100105, to the Department of Computer Science and Engineering of SRM University Delhi-NCR, Sonipat, Haryana, India in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering (specialization in Data Science and Artificial Intelligence) under the faculty of Engineering and Technology is an authentic record of the work carried out by her/him under my supervision. In my opinion, this work fulfills the requirement for which it has been submitted.

This project has not been submitted to any other University or Institution for any other degree and is submitted as a major project (Course Code CS4114) in 8th semester during the academic year 2022-2023.

Dr. Puneet Goswami

(Supervisor)

H.O.D. Dept. of Computer Science and Engineering

Ms. Lakshita Aggarwal

(Co-Supervisor)

Asst. Professor Dept. Of Computer Science

ACKNOWLEDGEMENT

I would like to express my gratitude towards staff of Computer Science and Engineering Department as well as the vice-chancellor of SRM University Delhi-NCR, Sonipat for providing me the opportunity to complete a project on an institutional repository system.

My sincere thanks to Dr. Puneet Goswami and Ms. Lakshita Aggarwal for their support and guidance for completion of this project.

CANDIDATE'S DECLARATION

I hereby certify that the work being presented in the project entitled "**INSTITUTIONAL REPOSITORY SYSTEM**" in partial fulfilment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering of SRM University Delhi-NCR, Sonipat, Haryana, India is an authentic record of my own work carried out under the supervision of Dr. Puneet Goswami and co-supervision of Ms. Lakshita Aggarwal, as major project in 8th semester during the academic year 2022-2023. The matter presented in this project has not been submitted for the award of any other degree of this or any other Institute/ University.

(Signature of the candidate)

Ayush Dobhal

(11019210066)

(Signature of the candidate)

Harsh Gupta

(11019210081)

(Signature of the candidate)

Sumeet Jain

(11019210105)

TABLE OF CONTENTS

ABSTRACT

In the modern era, the age-old practice of using physical documents has become quite inconvenient. This motivated the usage of digital documents for a myriad of purposes. While digital copies of documents can be a potential solution to issues of inefficiency in information management, proving their authenticity can make document verification problematic.

Document verification is a convoluted task involving several challenges and time-consuming authentication processes. Furthermore, distinct types of documents give rise to the need for customized authentication and verification procedures. The lack of an anti-forge mechanism makes the forgery of documents an easy task. For students in particular, academic certificates issued by their universities are of utmost importance. When applying at institutions or companies, students usually have to provide these certificates. Due to low transparency in the issuing process, skillfully generated counterfeits become hard to detect and verify. This stark increase in forged documents not only questions the credibility of the document holder but also jeopardizes faith in issuing authority. Furthermore, corporations annually invest heavily in job applicants' background verification process. Yet the verification process implemented by these companies to determine the authenticity of academic certificates is quite pricey, time-consuming, and inefficient. The task of keeping tract of every certificate and manually confirming its legitimacy has become tiresome. Thus, in compliance with the principle of confidentiality, reliability, and availability, academic certificates need to be digitalized.

Addressing the issue of counterfeiting certificates, we have proposed a digital certificate verification system based on principles of blockchain technology. This verification system can effectively solve the current issue of confirming the legitimacy of digital documents, for example, an image of a birth declaration, a marked authoritative report determining an agreement, etc. at an exceptionally low implementation cost. Once used effectively, it has the potential to emerge as a significant tool to combat document fraud and misuse. Moreover, anybody with access to the platform can now check the authenticity of a document without depending on third-party verification. Also, students face less risk of losing or damaging a certificate, making their data more secure and safe. To the best of our knowledge, our aim is to identify the gaps and loopholes in the current blockchain-based academic certificate verification solutions and to overcome them by creating a novel framework for an Institutional Repository.

Keywords— Blockchain, document verification, institutional repository, digital certificate, authentication.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PKI	Public Key Infrastructure
IRES	Institutional Repository System
HTML	Hyper Text Markup Language
CSS	Cascading Style Sheets
API	Application Program Interface
PDF	Portable Document Format
CSV	JavaScript Object Notation
ID	Identity Document
DAPP	Decentralized Application
IPFS	Inter Planetary File System
CV	Curriculum Vitae
SHA-256	Secure Hashing Algorithm 256
SHA-2	Secure Hash Algorithm 2

Chapter 1
INTRODUCTION

1.1 GENERAL INTRODUCTION

Official figures state that India has a total of 1057 universities as per the University Grants Commission report [1]. With the much-needed development in the number of institutions, new obstacles have arisen in the current education system. One of these issues happens to be the credibility of academic certificates provided by educational institutions. Throughout the cycle of a student's study years from kindergarten to college, he/she needs to provide these certificates in each stage for validation. Academic certificates are an indicator of an individual's knowledge, competencies, and aptitude achieved via education. Academic documents are particularly important in situations involving employment and higher studies as they serve as proof of skills, knowledge, and expertise. Only academic qualifications granted by a university having legal authority to do so are considered to be authentic.

With the rapid digitization in every sector, the usage of physical documents has become quite inconvenient. Moreover, there is a potential risk of losing and damaging all these important certificates. This has given rise to the need for digital certificates. Due to the invaluable nature of these certificates, unqualified people tend to lie about their academic background by producing illegitimate certificates. Forgery of graduation certificates has become a common malpractice that not only tarnishes the university's reputation but also has an impact on the hiring in government, public and private sectors. Fabricating documents is not limited to generating a fictitious degree or certificate from a recognized institution but also involves modification in legal documents that happen to be alterations in legitimate documents such as changes in grades, date of birth, course specialization, enrollment dates, graduation dates, etc. Moreover, the lack of a normal convention to validate physical/digitized documents makes analyzing records troublesome. In India, there are 21 fake universities operating and myriad fake degree certificates in circulation [2]. Thus, the lack of an efficient anti-forge mechanism makes the forgery of documents an easy task.

Multiple observations suggest that a majority of candidates lie at least about some part of their educational credentials and professional experience. [3]states that there are three primary channels via which fake credentials are generated: systematic fraud, organized fraud and independent fraud. The ever-growing cost of higher education and the high rate of unemployment in India fuels the demand for forged certificates in students and job seekers. In the hiring process, educational details are of crucial importance. [4] highlights the need for applicant's education verification in order to avoid negligent hiring liability. If an organization fails to verify the background details of hired employees, it could be held liable for negligent hiring. Due to prevalence of counterfeit documents provided by the job seekers, the hiring process has become a tedious undertaking for the human resource departments of organizations. In India, due to the absence of an efficient document verification application, companies end up hiring some third-party agencies to do background verification of all the newly hired employees and end up spending a lot of money on this process.

Validating the certificates properly before taking someone into an organization is the only way to solve this problem. The root cause of this issue of forgery is that credential verification is not as easy as it is seeming. Tracking all certificates and validating their authenticity manually becomes a tedious task. While the perpetrators of forgery might be punished or pardoned yet the monetary loss faced by organizations and institutions can be hefty.

University students require authentic certificates to pursue studies or to apply for jobs. These credentials provided by millions of graduating students are verified via a lengthy document verification process Yet due to uncertain circumstances, they often lose their educational certificates. Furthermore, reapplying for hard copies can be tedious and time-consuming as certifications are granted by different organizations and in-person application may be a necessity. This results in significant overhead as documents are transferred between the institution and the student multiple times. In contrast to this age-old adage, applying for an e-copy can save time and resources. By providing information for identity verification, graduates can access these documents. Moreover, recruiters can identify the genuineness of these documents on the spot without questioning the credibility of issuing authority thereby saving the need for background checks. An automated identity verification system, entirely maintained by the university will ensure that the data is genuine.

The rapid development in information technology makes protecting sensitive data a necessity. In today's world, most systems are susceptible to massive data breaches. These can lead to unauthorized people gaining access to sensitive documents, thereby contributing to potential identity theft and increased accessibility to counterfeit documents.

Blockchain emerges as a potentially effective solution to the issues mentioned above. It serves as an incontestable solution to the problem of document verification as it provides us with numerous benefits such as security, authentication and integrity as all transactions are safely recorded. It provides a trusted environment where actions committed are visible and cannot be tampered with. Emulating the fundamental principles of blockchain, an Institutional Repository can be made to so as cater to the rising need for a verified and tamper proof credential verification system.

An institutional repository is a database to store and distribute a university's scholarly research in digital formats. On a large scale, they can act as digital collections that capture and preserve the intellectual output of a single or multi-university community [5]. While conventional institutional repositories primarily aim at providing access to scholarly material, they have a promising potential in being of much use to researchers. If used at its full potential, it can act as an efficient identity verification system that not only preserves digital documents but also keeps their authenticity intact. Convenience, accessibility, satisfaction, and usefulness are the determining factors in the use of such electronic databases. In today's world, we are extensively using the PKI (Public key infrastructure) model which does not guarantee the authenticity of digital documents. Since the internet system is not designed for this protocol to verify anything digitally, we intend to use our application: Institutional Repository System (IRES) to keep the data secure and tamper-free. Just like in blockchain, we use the SHA-256 algorithm to encrypt or secure the document.

This proposed project finds its usage in providing a secure Digital tamperproof record by universities. It can be utilized for providing Scholar and Medical Certificates.

Besides that, it can be used as a tool to verify the documents of employees by organizations. The platform will be accessible to university authorities, Students, and Certificate Verifiers. The activities that can be performed by each party are as per the following:

Universities:

1) Issue academic certificates.
2) View academic certificates issued.
3) Endorse Verification and digitally sign academic certificates.

Students:

1) Receive academic certificates from universities.
2) View and manage received academic certificates.
3) Share academic certificates with third-party verifiers.
4) Selective disclosure of certificate data.

Verifier:

1) Receive certificate data from students.
2) Verify certificate authenticity with the IRES platform.

1.2 APPROACH TO THE PROBLEM IN TERMS OF TECHNOLOGY

The problem of document verification will be solved using blockchain technology. The following tech stack will be used for developing the application:

- **HTML-** Hyper Text Markup Language, also known as HTML, is used for building websites and online applications. It is used for making visually appealing web pages with the help of styling. HTML documents are composed of multiple HTML tags, each tag containing different content.

- **CSS-** Cascading Style Sheets is the language for creating style sheets that describe the layout and appearance of documents written in markup language. It is usually used with HTML to modify the web page style and user interfaces. Typically, CSS is used with HTML and JavaScript in a majority of websites to develop user interfaces for web as well as mobile applications.

- **JavaScript-** It is a dynamic computer programming language that is most commonly used as part of web pages. The implementation of JavaScript allows the client-side script to interact with the user thereby helping in the creation of dynamic pages. Prior, it was termed LiveScript.

- **ReactJS-** It is a popular JavaScript front-end library for creating reusable UI components. It is an open-source, component-based library that is responsible for the view layer of an application

- **Node.js-** Developed by Ryan Dhal, Node.js is an open-source, cross-platform built on Google Chrome's JavaScript Engine (V8Engine). It uses an event-driven, non-blocking I/O model, perfect for data-intensive real-time applications.

- **Firebase**- An app development platform backed by Google for building mobile and web applications. It aids developers build their applications fast and in a secure manner. NoSQL database is used to store the data.

Chapter 2
LITERATURE SURVEY

2.1 SUMMARY OF THE PAPERS STUDIED

Lakmal et al. [6]

IDStack - The Common Protocol for Document Verification built on Digital Signatures (2019)

Non-transparency in the verification process is a big issue when using soft copy of documents. These documents are often susceptible to deliberate modifications and unintentional errors. To resolve this issue of document verification, the authors of the paper [6] proposed IDStack, a framework that utilizes digital signatures, text extraction, and document correlation as a solution. It is a decentralized API stack that can be accessed by users through a client application that calls the IDStack protocol. IDStack architecture is composed of three modules which are data extraction, data validating, and score calculating module. These modules focus on three key applications that are connected to the IDStack web service. Modules can be accessed by the extractor, validator, and relying party. The extractor is the owner or any third-party user that can authenticate and sign digitally stating that the physical document and the digital document are matching. A validator is a third-party user that can attest as well as digitally sign stating that the content is valid, the preceding signature is valid, or both are valid. The relying party is the party that wants to view the documents. Citizens have the liberty to play any role in this protocol.

This digital signature method assigns confidence score and a correlation score to documents. Information from a document is extracted to determine the confidence score. Citizens with physical or digital copies of documents ask an extractor to verify and digitize them. The document is digitized by the extractor, who uses IDStack to build a machine-readable document that certifies that the physical or digital document and the machine-readable document are a perfect match. This document is then certified by the digital signature of the extractor. Docparser is a service that translates PDF files to machine-readable formats like JSON, CSV, etc. The intended protocol works similarly to how machine learning aids in data extraction by automatically detecting fields like name and address in pdf documents. With the use of an automatic real-time scoring mechanism that determines a score based on the signatures and the correlativity of the document parameters, the relying party can now access these signed machine-readable documents.

There are some limitations to this solution. The entire process takes a long time. Since various papers may have different languages, the fact that IDStack only allows English documents can be problematic. Calculating confidence scores is quite difficult with blurry documents. Additionally, the suggested method relies on self-signed certificates for digital signing, allowing users to generate their own signatures. Also, there is a short-term fix that improves the system's usability, but it introduces the challenge of confirming the identity of the signer. Therefore, we can compute a calculative score based on how frequently that document is correctly validated from a decentralized chain rather than computing confidence score from data extraction.

Reddy et al. [7]

Proposing a reliable method of securing and verifying the credentials of graduates through blockchain (2021)

The research paper [7] states that there is a lack of a secure platform to store academic certificates and verify them as per the requirement. Although some universities digitally store certificates in their centralized network, these certificates are vulnerable to tampering. This issue primarily arises due to a lack of timestamp and a secure way to store data.

As keeping the stored certificates secure is a matter of utmost importance for the university and students, the authors propose a system to store and verify student credentials using blockchain technology. On adding a certificate into a block, the system will return an exclusive certificate ID and Aadhar card number of the student as a primary key. By using the unique certificate ID, it can be verified whether the certificate is authentic or not. By using the student's unique Aadhar card number, the verifier can have a look at all the authentic certificates of that student.

While adding the certificate, the certificate authority needs to compensate with some Ethereum gas value which shall be debited from the account of the certificate authority. This Ethereum gas value is paid to miners for adding blocks to the blockchain. The distributed nature of blockchain makes it hard to tamper with the data stored in a block. For example, a blockchain is created having some certificates as blocks of data over a period of time. In a scenario, an intruder tries to alter marks in a specific block, the hash value will not remain the same, thereby reflecting the tampering done in succeeding blocks of the blockchain. If the intruder somehow alters the next block in a particular node, the change cannot be propagated to all other nodes due to their distributed nature. Thereby, if any changes are made, they become easily identifiable during the consensus process. This removes the threat of altering student grades which is a big issue in existing systems where student marks are stored in a central database of a university server. Another advantage of this application is that employers can verify documents easily in no time. It overcomes the disadvantages of the traditional credential verification phase wherein there is always a possibility of manipulating the concerned section authority to manage the verification process.

The system is implemented and tested using: JavaScript, Truffle, Solidity, Ganache, Ethereum, and MetaMask. Ganache, which happens to be a part of the Truffle ecosystem is utilized for developing the Decentralized Application. Once the DAPP is developed and tested on ganache, it is ready to be deployed on the Ethereum client. Truffle essentially helps in developing, testing, and deploying the DAPP. MetaMask, a digital currency wallet is used for transactions using Ethereum-based tokens.

Using this system, academic credentials can be authenticated using either the transaction ID of the certificate or the Aadhar number of the student. This system can be used by universities to keep students' data secure. The future scope of this system involves its utilization in proving the validity of any type of document belonging to a myriad of sectors wherever there is a requirement for a digital document timestamp. Upon further work, the authors intend to use the system for storing, tracking, and verifying employment and experience data of students and job applicants.

Jamal et al. [8]

Blockchain-Based Identity Verification System (2019)

This research paper [8] suggests a solution that keeps a user's private data on the blockchain. Users can be categorized as individuals, authorities, or third parties. These users' data is kept secure because of blockchain's security features. A person can give permission for a third party to view their records and confirm their legitimacy. While this is going on, the authority can verify organizations that want to be registered requesters and upload users' records to the blockchain. Users' data may be requested by third parties for verification purposes. Thus, the system serves as a safe repository for a person's personal data that others can use for identity verification, maintaining validity and eliminating the need for further verification.

The development process known as Agile Unified Process was used to ensure that the project would continue to advance until the successful conclusion of the final phase. In the early phase, the project's fundamental components are established together with its scope and limitations. In the second phase, the system's design was developed using diagrams created using the Unified Modelling Language. In the third phase, the system was built using Microsoft Visual Studio Code and Android Studio 3. System and acceptance testing were part of the final phase. The Huawei Mate 8 was utilized as the test device because the system was created to be both a web and a mobile application.

Some of the functionality offered by the system includes features for uploading and requesting user details, logging in for users and authorities, and user registration. A survey with three questions was used to gather the results of the acceptability testing.

The project should be carried out on a real working blockchain rather than just a local system, according to the future scope of improvement, to enable system commercialization. Additionally, voice recognition and fingerprints can be used in future generations of biometrics to lock and unlock the application. Additionally, the ability to keep a range of personal records might be added since the existing system only allows for the storage of users' passport information. Last but not least, screenshots can be restricted to maintain their confidentiality.

Lamkoti et al. [9]

Certificate Verification using Blockchain and Generation of Transcript (2021)

Building an immutable certificate generation and validation system is the main goal of the paper [9]. This proposed solution provides a simple method for establishing a document's legitimacy. Furthermore, using Ethereum smart contracts, it is possible to assess the authenticity and originality of these documents. The college serves as the issuing body for certificates under this system. It is permitted for students to download and view their corresponding digital certificates. They are also given a hash value linked to the document so they can access and check documents in the future. Businesses will be able to check the validity of academic credentials.

The Certificate Issuing Authority will feed the Certificate Template into the system. The user, who is the student, will now enter the information needed to produce the academic certificate. An electronic document preview will result from this. The information will be added to the provided certificate template after validation. A certificate preview will result from this. The user's confirmed information will be stored when the user has reviewed and approved the document. This ensures that the hash value stored in the blockchain cannot be altered normally. This hash or the digital certificate may now be sent by the student to the desired entity. On the other side, the issuer can enter the hash key or upload this certificate. The system will subsequently produce a response indicating whether the certificate is genuine or not.

Solidity, a high-level language derived from JavaScript, was used to create the code for this application. The Migration. Sol and Certification. Sol contracts make up the majority of the code. The default Solidity file included with the Ethereum setup is called Migration. Sol. It maintains track of all transactions and contains code dictating how they should proceed. The system's primary code is located in the Certification. Sol file. Additionally, it includes a struct Certificate that will serve as the data block's structure. This is a guide for how the information will be stored in the Ethereum block. The uploaded documents are done so using the Inter Planetary File System. The distributed nature means that the submitted papers are safe and only available with the help of hash. Immutable ledgers aid in creating a system where all transactions are unchangeable and transparent. This solution decreases the amount of manual effort required for verification while also automating the process of creating certificates. Students are less likely to lose their certifications as a result. The likelihood of data manipulation is greatly diminished as a result of the additional hashing algorithm. While the actual certificate will be kept in IPFS, the blockchain will only store the certificate's hash.

Marella and Vijayan et al. [10]

Document Verification using Blockchain for Trusted CV Information (2020)

The authors of the paper [10] suggest a viable solution to verify the background details of job applicants using blockchain technology. Design Science Research Methodology approach was used in order to develop an Information System for solving the problem.

On careful examination of existing background verification processes, three key objectives were determined: the solution must flawlessly conduct background checks on job applicants and be cost-efficient and time-saving.

After careful evaluation, Hyperledger fabric was used for developing the solution. A consortium blockchain is made where the privilege to write information onto the blockchain is given to universities, companies, police, doctors, and certificate authorities. Universities will have the provision to submit all the academic details of students. Companies can fill in the details of their employees regarding work experience on the blockchain. Doctors can submit the medical reports of the applicants. Police will have the provision to submit documents that verify the job applicant’s criminal history. Finally, the certification authority can issue certificates to users upon successful completion of training. An administrator node shall be responsible for verifying all these documents before the calculation of hash values and then saving them to the blockchain.

The transaction should be verified and approved by the organization's administrator. On approving the transaction, the hash value of the document involved in the transaction and the hash value of identification are calculated and provided to the consensus mechanism for updating the blockchain. On the blockchain, the document's hash value and the identification hash are both stored. The hash values become unchangeable and impossible to manipulate once they are stored on the blockchain. During the employment process, the hiring manager will upload the applicant's supporting documents to the hash verifier web application that runs on the blockchain. The uploaded document's hash value will be compared to the hash value stored on the blockchain by the hash verifier application. The application returns a value as legitimate if both hash values are equal. If not, a value indicating that the document is a fake is returned. As a result, the hiring manager can quickly determine whether the documents submitted by the candidate are genuine.

The hiring manager only needs to pay a small amount of money according to the proposed solution for the hash verifier application to authenticate each document. The organization entities that authenticated the transaction and saved the document's hash value to the blockchain will be credited with this money. During the recruiting process, this low-cost background check procedure will assist small and medium-sized businesses in efficiently verifying an applicant's background information. Moreover, hiring managers can instantaneously verify the applicant's documents during the hiring process itself. Once an organized ecosystem is built, the system can be used globally as a unified solution for verifying the CV provided by job applicants.

2.2 TABULAR COMPARISON

Table 2.1 Tabular Comparison

Reference Number	Title	Year of Publication	Feature	Advantage
[6]	Document verification using digital signatures.	2019	Makes the process of document verification time consuming yet simple and secure.	Provides an effective framework that utilizes digital signatures, text extraction, and document correlation as a solution for information verification
[7]	Blockchain as a reliable method of securing and verifying the credentials of graduates.	2021	Using this system, academic credentials can be authenticated using either the certificate's unique transaction ID or the student's Aadhar number.	Upon further work, the system can be utilized in proving the validity of any type of document belonging to a myriad of sectors wherever there is a requirement for a digital document timestamp.
[8]	Blockchain-Based Identity Verification System	2019	The system to verify the user's identity is available as a web application and a mobile application.	Voice recognition and fingerprints can be used in future generations of biometrics to lock and unlock the application.
[9]	Certificate Verification using Blockchain and Generation of Transcript	2021	This solution decreases the amount of manual effort required for verification while also automating the process of creating certificates.	The Certificate Issuing Authority can feed a Certificate Template into the system. Thus, colleges can utilize this system to issue digital certificates.
[10]	Document Verification using Blockchain for Trusted CV Information	2020	The system allows institutes, corporations, police, doctors, and certificate authorities to write information of job applicants onto the blockchain which recruiters can easily verify.	Hiring managers can instantaneously verify the applicant's documents during the hiring process itself using this proposed solution.

2.3 INTEGRATED SUMMARY OF THE LITERATURE STUDIED

For the literature study, we took 5 research papers in which blockchain technology was being used for skill verification of digital documents. Our findings from each research paper are as follows:

- **Lakmal et al. [4]-** The authors of this paper proposed IDStack, a framework that utilizes digital signatures, text extraction, and document correlation as a solution to issue of document verification. It is a decentralized API stack that can be accessed by users through a client application that calls the IDStack protocol. This digital signature method assigns confidence score and a correlation score to documents. Information from a document is extracted to determine the confidence score. Citizens with physical or digital copies of documents ask an extractor to verify and digitize them. The document is digitized by the extractor, who uses IDStack to build a machine-readable document that certifies that the physical or digital document and the machine-readable document are a perfect match. This document is then certified by the digital signature of the extractor. With the use of an automatic real-time scoring mechanism that determines a score based on the signatures and the correlativity of the document parameters, the relying party can now access these signed machine-readable documents. The entire process is time-consuming. Also, calculating the confidence score of blurry documents is difficult.

- **Reddy et al. [5]-** The main goal of the paper is to build an immutable certificate generation and validation system. This proposed solution provides a simple method for establishing a document's legitimacy. Moreover, using Ethereum smart contracts, it is possible to assess the authenticity and originality of these documents. The college serves as the issuing body for certificates under this system. It is permitted for students to download and view their corresponding digital certificates. They are also given a hash value linked to the document so they can access and check documents in the future. Businesses will be able to check the validity of academic credentials The Certificate Template is put on the system by the Certificate Issuing Authority. The user, who is the student, will now enter the information needed to produce the academic certificate. Once the student verifies the details,

- **Jamal et al. [6]-** This research paper suggests a solution that keeps a user's private data on the blockchain. Users can be categorized as individuals, authorities, or third parties. These users' data is kept secure because of blockchain's security features. A person can give permission for a third party to view their records and confirm their legitimacy. While this is going on, the authority can verify organizations that want to be registered requesters and upload users' records to the blockchain. Users' data may be requested by third parties for verification purposes. Thus, the system serves as a safe repository for a person's personal data that others can use for identity verification, maintaining validity and eliminating the need for further verification. Some of the functionality offered by the system includes features for uploading and requesting user details, logging in for users and authorities, and user registration. A survey with three questions was used to gather the results of the acceptability testing.

- **Lamkoti et al. [7]-** The authors of this paper intend to build an immutable certificate generation and validation system. The proposed solution provides a simple method for establishing a document's legitimacy. Furthermore, using Ethereum smart contracts, it is possible to assess the authenticity and originality of these documents. The college serves as the issuing body for certificates under this system. It is permitted for students to download and view their corresponding digital certificates. They are also given a hash value linked to the document so they can access and check documents in the future. Businesses will be able to check the validity of academic credentials. This solution decreases the amount of manual effort required for verification while also automating the process of creating certificates. Students are less likely to lose their certifications as a result. The likelihood of data manipulation is greatly diminished as a result of the additional hashing algorithm. While the actual certificate will be kept in IPFS, the blockchain will only store the certificate's hash.

- **Marella and Vijayan et al. [8]-** The authors of the paper propose a solution for background verification of job applicants using blockchain technology. Design Science Research Methodology approach was used in order to develop an Information System for solving the problem. Hyperledger fabric was used for developing the solution. A consortium blockchain is created which grants universities, businesses, medical professionals, law enforcement and certificate authorities the right to add information to the blockchain. During the recruiting process, this low-cost background check procedure will assist small and medium-sized businesses in efficiently verifying an applicant's background information. Moreover, hiring department can instantaneously check authenticity of the applicant's documents while the hiring process is ongoing. Once a well-organized system is developed, it could be utilized internationally as a primary solution to information verification of CV provided by job.

2.4 PROBLEM STATEMENT

A large number of counterfeit certificates in circulation has been a long-standing problem. The granting of such certifications has evolved into a business driven by people’s need for work. Hardworking persons with legitimate degrees/certificates must pay the consequences of this issue since those with falsified credentials deny them what may have been theirs. In many cases, this can be extremely detrimental. The project is an attempt to overcome this problem and for this, we will use the conceptual principles of blockchain to make a framework for Institutional Repository System. It will recognize the documents and display them whether the documents are issued by the authority or it is forged.

Chapter 3

PROPOSED METHODOLOGY

3.1 OVERALL DESCRIPTION OF THE PROJECT

This project intends to build a framework for institutional repository has the potential to revolutionize the academic credential verification process. It will allow for a more streamlined and efficient method of verifying academic credentials, reducing the time and resources required to complete the verification process. This will be particularly beneficial for employers who receive many job applications and need to verify the academic credentials of each candidate.

Moreover, the platform will be able to provide a higher level of security compared to traditional methods of information verification. The use of hashing can ensure the integrity and immutability of data stored on the platform, making it virtually impossible for anyone to tamper with the information. This will greatly reduce the risk of fraudulent activity and increase the trust between universities, employers, and students. The institutional repository can also serve as a valuable resource for students and faculty members. It will allow students to access their academic records and certificates anytime, anywhere. This can be particularly beneficial for students who have lost their certificates or need to access their records for further studies or job applications. Faculty members can also use the platform to store and share their research publications, creating a collaborative space for knowledge sharing. Overall, the proposed institutional repository framework has the potential to revolutionize the way academic credentials are verified and stored. Its implementation can bring about a much-needed change in the academic and employment sectors, improving productivity, security, and efficiency.

This project will be considered successful when the User Interface and functionality is successfully developed for all three categories of users of the platform and authentication can executed.

3.2 FUNCTIONAL REQUIREMENTS/ NON-FUNCTIONAL REQUIREMENTS

3.2.1 Functional Requirements: -

- **React JS-** It is a popular JavaScript front-end library for creating reusable UI components. It is an open-source, component-based library that is responsible for the view layer of an application.

- **Hashing Algorithm-** A cryptographic hashing technique (or function) is used to verify the integrity of messages, files, and data. It belongs to the SHA-2 family of hash functions and employs a 256-bit key to transform a piece of data into a new, unrecognizable data string of a defined length.

- **Material UI** - Material UI is a free and open-source React component library based on Google's Material Design. It comes with many prebuilt components that are ready for use in production straight away. Material UI is designed to be appealing and has a range of customization tools that make it simple to build your own customized design system on top of our components.

- **Cloud Technology-** The concept of "Cloud technology" describes the process of storing, managing, and processing data on remote servers connected to the internet. With the ability to access data and applications from any device and place with an internet connection, the cloud offers both organizations and consumers flexibility and scalability. In this Firebase is a popular choice for developing mobile and web applications due to its ease of use, flexibility, and integration with other Google services.

3.2.2 Non-Functional Requirements: -

- **VS Code-** Visual Studio Code is a code editor by Microsoft for operating systems such as Windows, Linux and macOS. It provides a multitude of features like embedded Git, support for debugging, code refactoring.

- **NPM -** It is an online repository for the publishing of open source Node.js projects. It is a command line utility for interacting with said repository that aids in package installation, version management and dependency management.

- **Browser-** A web browser is a program that allows you to access websites. When a user requests a web page from a certain website, the browser receives the page's files from a web server and displays it on the user's screen. Browsers are used on a variety of devices such as desktop computers, laptop computers, tablets, and smartphones.

3.3 BLOCK DIAGRAM

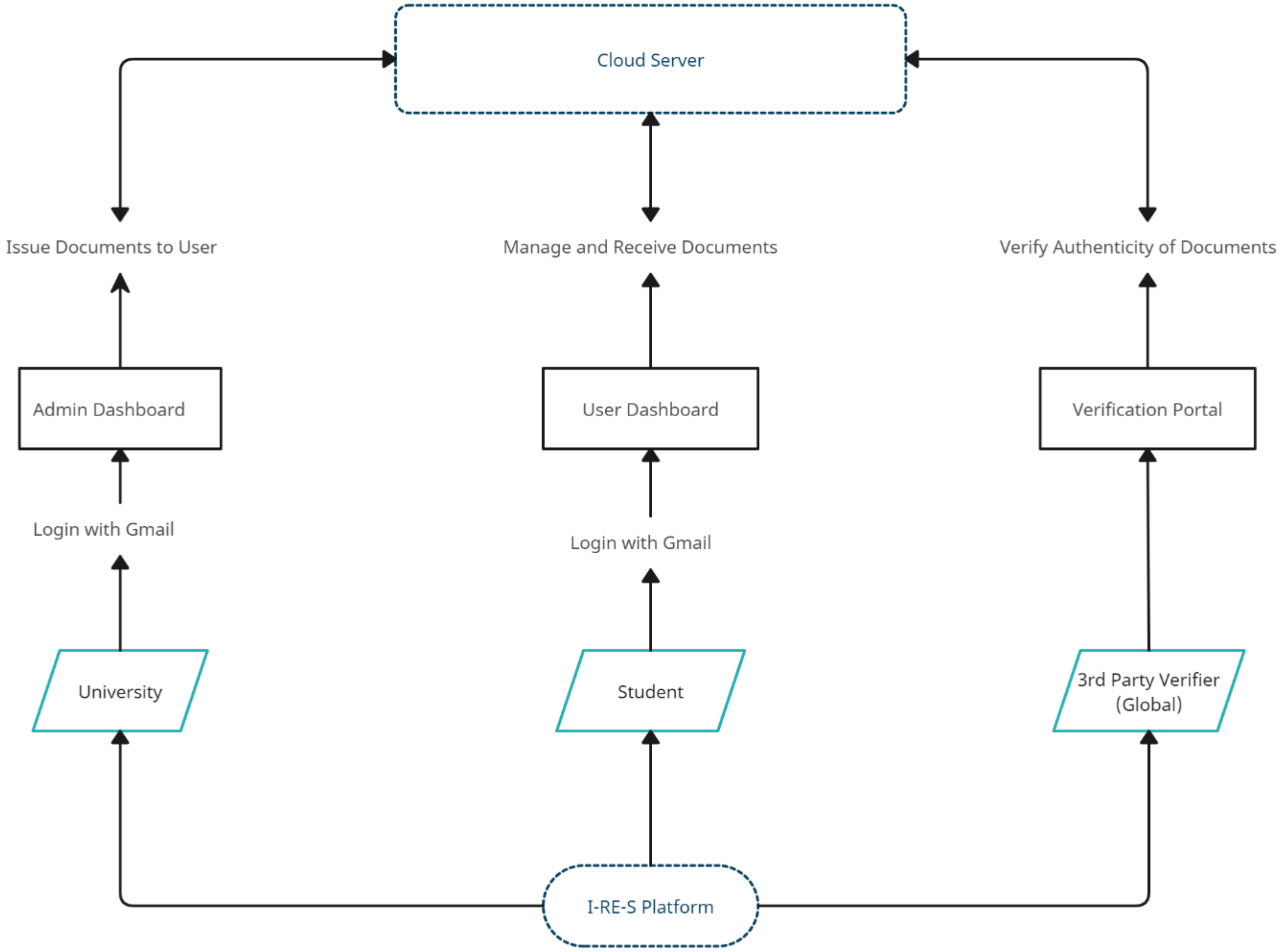


Figure 3.1: Architecture for Institutional Repository System

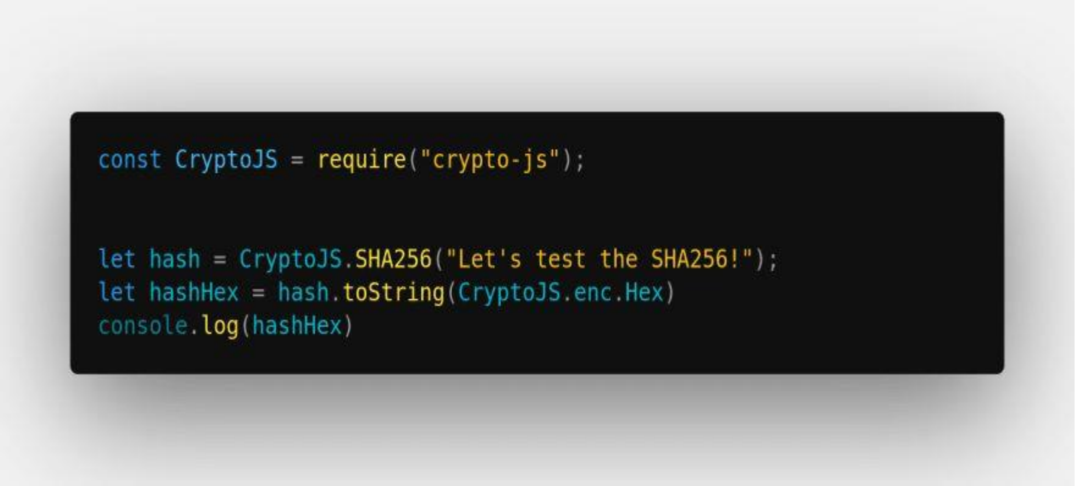
Figure 3.1 indicates an overall working of I-RE-S Platform. The platform is designed keeping in mind the objective of enhancing user experience by dividing platform users into three categories: University (Authorizer/admin), Student (User), third-party verifier (Global). Each category has its unique set of roles. Authorizer can verify the documents; User can put in the request to get their documents verified and the third-party verifier can check the status of documents whether document is verified or not.

3.4 ALGORITHMS

With rapid advancements in technology, cryptographic hash functions have proven to be potent at maintaining authenticity and integrity of data present in digital systems. National Security Agency developed secure hash algorithms and employed them as one-way functions so that they could be processed in one way. This makes their reverse engineering a laborious task, thereby maintaining authenticity of encrypted data. Secure Hash Algorithm 256-bit, or SHA-256, is a popular cryptographic hash algorithm. It is a member of the family of hash functions known as SHA-2 (Secure Hash Algorithm 2) which were released in 2001. Due to many vulnerabilities observed in prior developed algorithms, the demand for better hash functions drove the creation of SHA-256. Secure Hash Standard specifies SHA-256 as the 256-bit variant of SHA-2.

The working of SHA-256 algorithm involves following steps:

- Data is first converted to its binary equivalent.
- The binary data is divided into 512-bit sized blocks. If a block is smaller than that, bits of padding will be added to bring it to prerequired block size.
- They will be further divided into smaller blocks, each being 32 bits.
- 64 iterations of compression functions are performed wherein the generated hash values are rotated in a specified pattern and additional data is added.
- From the output of prior operations, new hash values are created.
- In the final round, the final 256-bit hash value produced is called the hash digest and is the end product of SHA-256.



```
const CryptoJS = require("crypto-js");


let hash = CryptoJS.SHA256("Let's test the SHA256!");
let hashHex = hash.toString(CryptoJS.enc.Hex)
console.log(hashHex)
```

Figure 3.2: Code snippet of SHA-256 algorithm

SHA-256 offers a myriad of advantages:

- Security: Due to its collision resistance, preimage resistance, and second preimage resistance qualities, SHA-256 offers a high level of security.

- Computational Efficiency: Its strong computational efficiency allows fast computations even when some device suffers from resource constrains.
- Standardization: SHA-256 is widely accepted and standardized due to its strong security and high computational efficiency. Thus, it is used in blockchain based technologies such as Bitcoin for verifying transactions.

Data is transformed into fixed-length, essentially irreversible hash values using SHA 256, which is why it is primarily used to validate the authenticity of data. SHA-256 is the industry standard used by multiple organizations worldwide.

Chapter 4

RESULTS AND DISCUSSION

4.1 IMPLEMENTATION DETAILS AND ISSUES

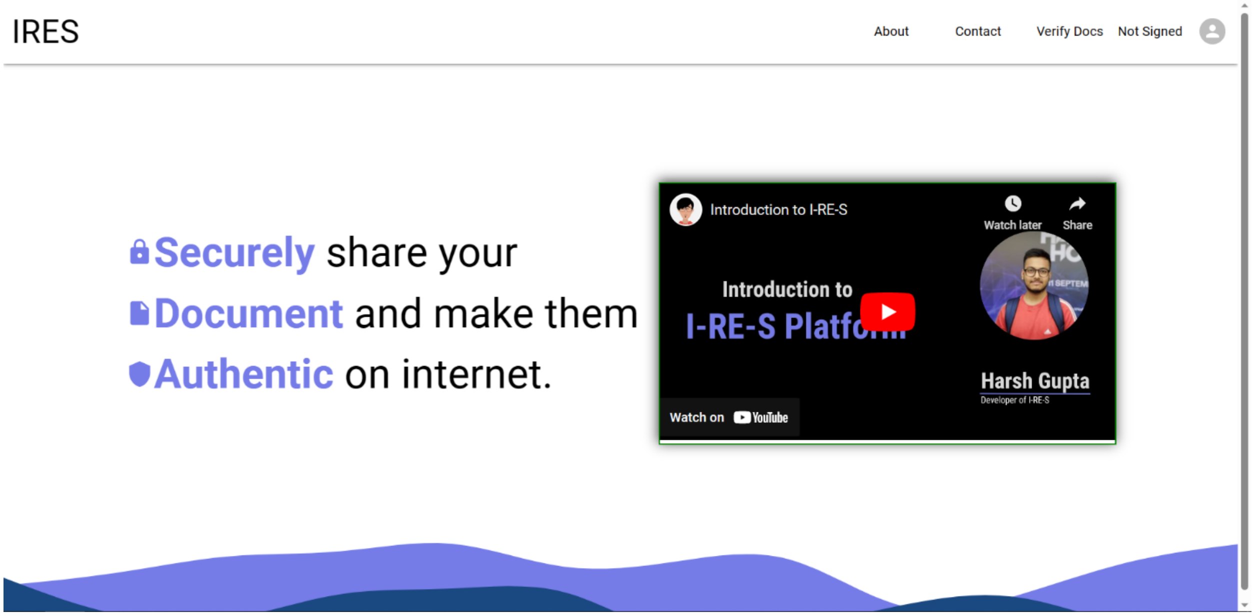
The implementation of this project primarily focused on developing the User Interface and functionality for all the users of our platform.

4.2 EVALUATION PARAMETERS

- Home Page should show menu bar with various options that can be used.
- Users should not face issues while signing or logging into their account.
- Navigation bar options should be functional and optimized.
- All Navigation Pages should be distinct and function correctly.

4.3 RESULTS

Following are the snapshots of the developed User Interface and Functionality:



Securely share your Document and make them Authentic on internet.

Figure 4.1 Home Page

Figure 4.1 presents the Home Page of IRES platform. It presents the login/ signup bar options and a video based brief introduction to the platform.

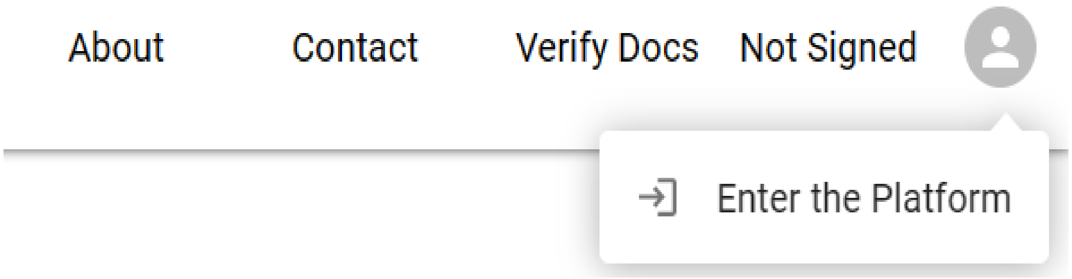
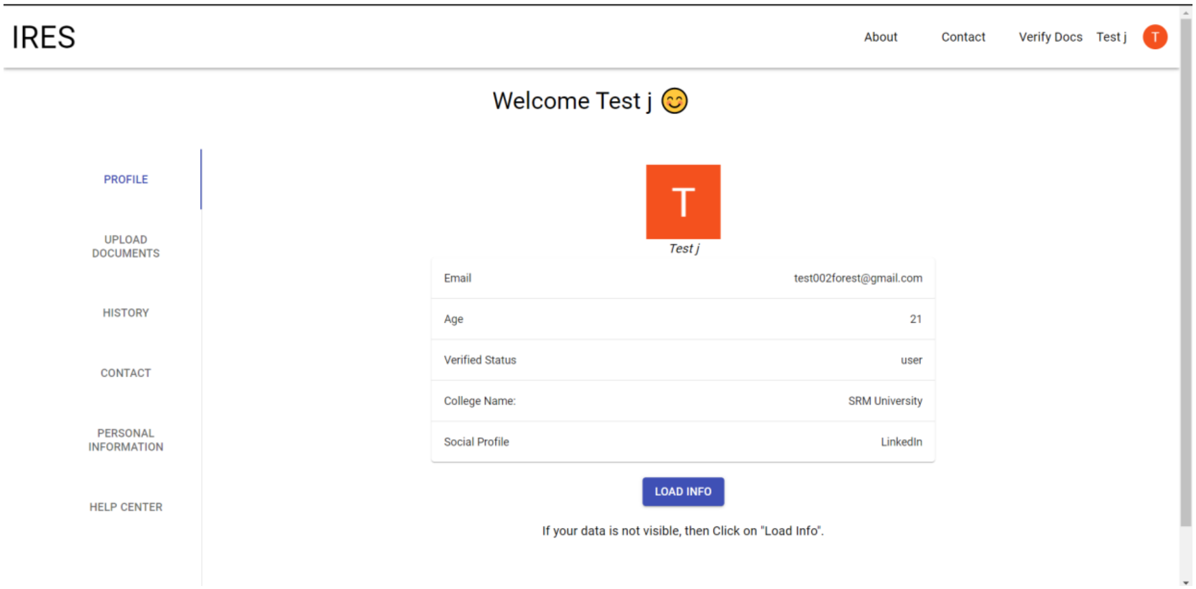


Figure 4.2 Login/ Signup Component

Figure 4.2 presents the options provided in Login/ Signup component. To use the platform, users have to sign up and log in with the help of Google Authentication.

Figure 4.3 Navigation Bar Options

Figure 4.3 depicts the Navigation Bar Options. To use the platform, users have to sign up and log in with the help of Google Authentication. The ‘About’ option provides the motivation behind developing this platform. ‘Contact’ option allows users of the platform to communicate their grievances, if any with the development team. User can proceed to view authenticity of a document using ‘Verify Docs’ option.



Welcome Test j 😊

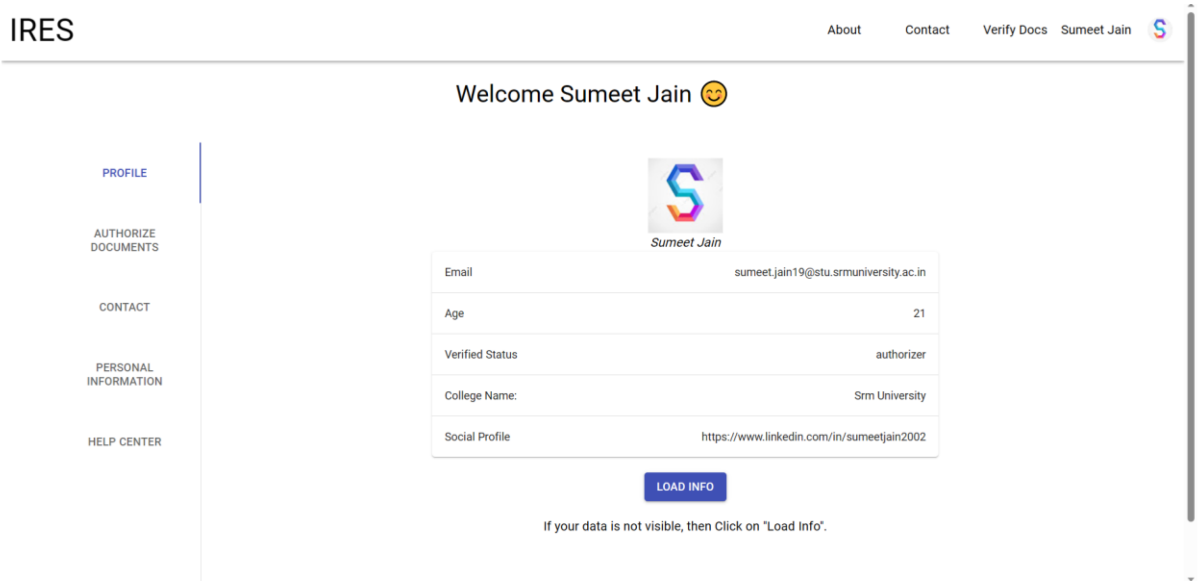
Test j

Email	test002forest@gmail.com
Age	21
Verified Status	user
College Name:	SRM University
Social Profile	LinkedIn

If your data is not visible, then Click on "Load Info".

Figure 4.4 User Profile Page

Figure 4.4 shows the User Profile Page. The recorded user metadata such as email, age, social profile etc. are displayed. 'LOAD INFO' option has been provided if in case user's data is not visible on the profile page. Other functionalities have been provided as options on left hand side of the profile page.



Welcome Sumeet Jain 😊

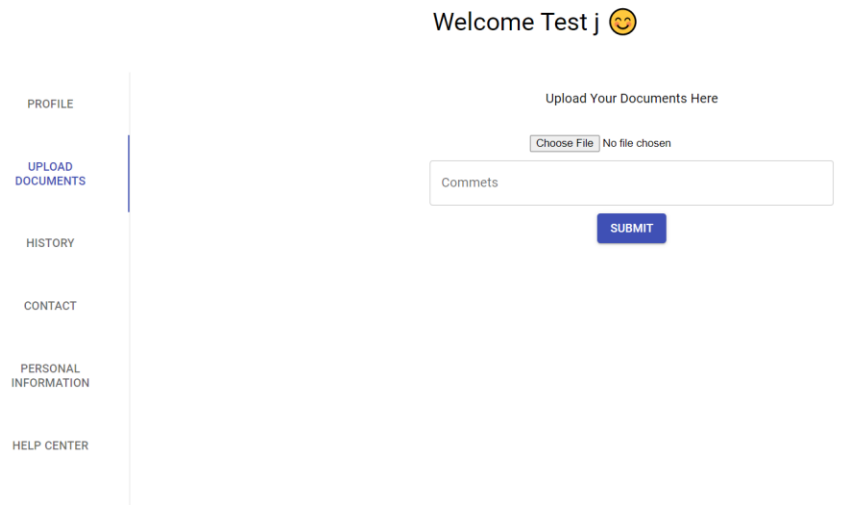
Sumeet Jain

Email	sumeet.jain19@stu.srmuniversity.ac.in
Age	21
Verified Status	authorizer
College Name:	Srm University
Social Profile	https://www.linkedin.com/in/sumeetjain2002

If your data is not visible, then Click on "Load Info".

Figure 4.5 Verifier Profile Page

Figure 4.5 presents the Verifier Profile Page. The recorded user metadata such as email, age, verified status etc. are displayed. 'LOAD INFO' option has been provided if in case user's data is not visible on the profile page. Other functionalities accessible only by the verifier have been provided as options on left hand side of the profile page.



Welcome Test j 😊

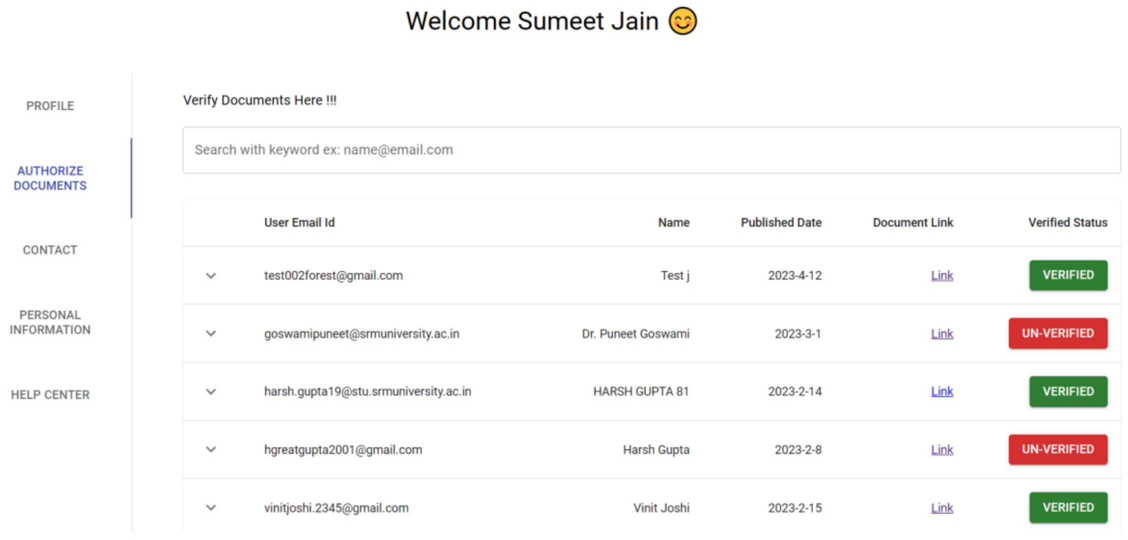
Upload Your Documents Here

No file chosen

Commets

Figure 4.6 User’s Document Upload Page

Figure 4.6 shows the User’s Document Upload Page. Here, User can upload their document in any format.



Welcome Sumeet Jain 😊

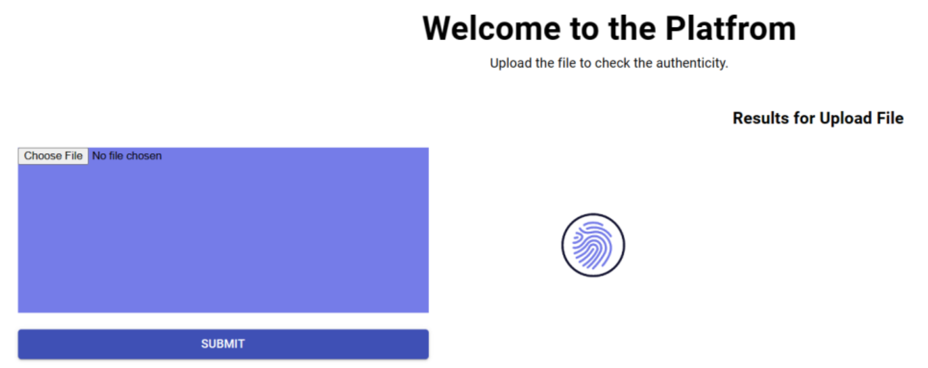
Verify Documents Here !!!

Search with keyword ex: name@email.com

User Email Id	Name	Published Date	Document Link	Verified Status
test002forest@gmail.com	Test j	2023-4-12	Link	VERIFIED
goswamipuneet@srmuniversity.ac.in	Dr. Puneet Goswami	2023-3-1	Link	UN-VERIFIED
harsh.gupta19@stu.srmuniversity.ac.in	HARSH GUPTA 81	2023-2-14	Link	VERIFIED
hgreatgupta2001@gmail.com	Harsh Gupta	2023-2-8	Link	UN-VERIFIED
vinitjoshi.2345@gmail.com	Vinit Joshi	2023-2-15	Link	VERIFIED

Figure 4.7 Admin Approval Page

Figure 4.7 presents the Admin Approval Page. Here the documents uploaded by users get verified.



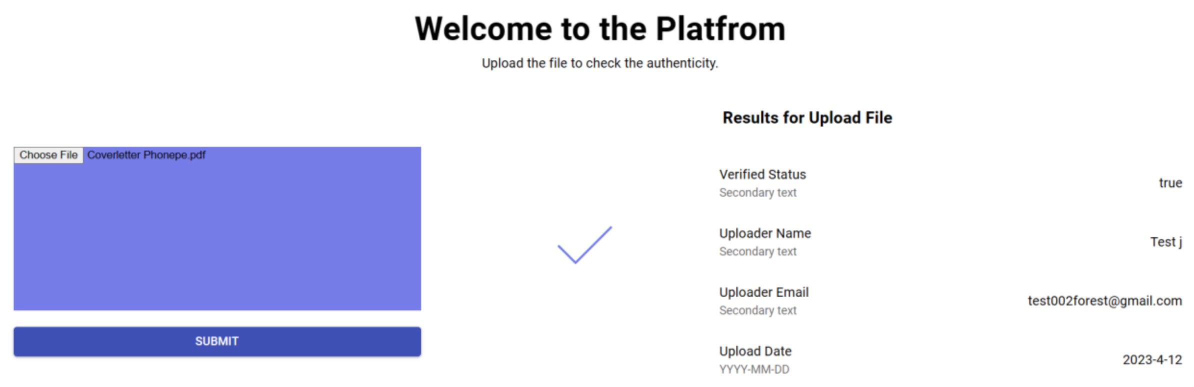
Welcome to the Platfrom

Upload the file to check the authenticity.

Results for Upload File

Figure 4.8 Verify Document Page

Figure 4.8 presents the Verify Document Page. The users of this platform can submit their documents to get their authenticity checked.



Welcome to the Platfrom

Upload the file to check the authenticity.

Results for Upload File

Choose File Coverletter Phonepe.pdf

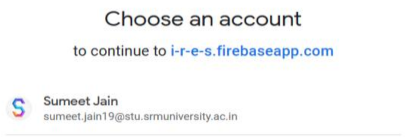
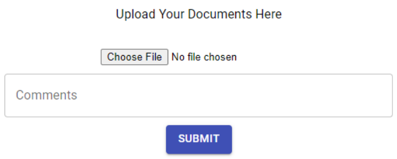
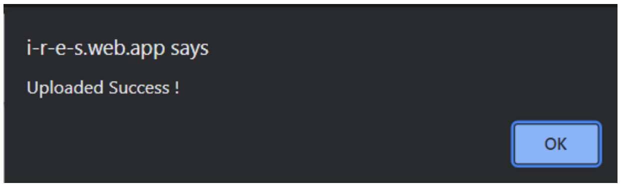
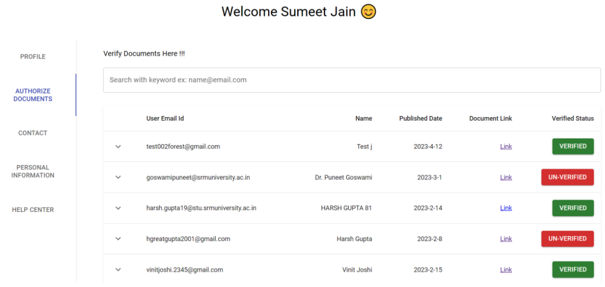
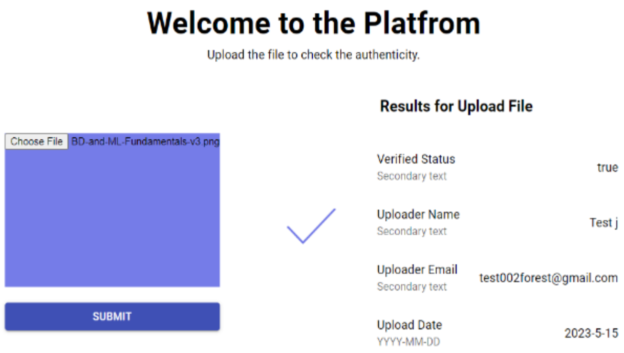
Verified Status Secondary text	true
Uploader Name Secondary text	Test j
Uploader Email Secondary text	test002forest@gmail.com
Upload Date YYYY-MM-DD	2023-4-12

Figure 4.9 Document Status Page

Figure 4.9 shows the Document Status Page. The verification status of the document or file uploaded by the user will be presented here.

4.4 OBJECTIVE MAPPING

Table 4.1 Objective Mapping

S. No.	Objectives	Parameters	Snapshot
1.	Record User metadata for upload documents like- social link, email.	Designed a sign up /sign in functionality by Gmail authentication so it can record all the information.	 Choose an account to continue to i-r-e-s.firebaseapp.com Sumeet Jain sumeet.jain19@stu.srmuniversity.ac.in
2.	Create digital fingerprint of documents by using hashing algorithm SHA-256.	User can upload their document in any format and on backend it will generate hash code and associate it with their profile metadata.	 Upload Your Documents Here Choose File No file chosen Comments SUBMIT
3.	Create verification request of their document to admin.	After uploading the document, script will make automatic request to admin, and that request will be added in their authorize document section.	 i-r-e-s.web.app says Uploaded Success ! OK
4.	Admin can preview verification documents to decide whether he should authorize them or not.	A section called 'Authorize Documents' where admin can track all request and verify from it.	 Welcome Sumeet Jain 😊 PROFILE AUTHORIZE DOCUMENTS CONTACT PERSONAL INFORMATION HELP CENTER Verify Documents Here !!! Search with keyword ex: name@email.com User Email Id | Name | Published Date | Document Link | Verified Status test002forest@gmail.com | Test j | 2023-4-12 | Link | VERIFIED goswamipuneet@srmuniversity.ac.in | Dr. Puneet Goswami | 2023-3-1 | Link | UN-VERIFIED harsh.gupta19@stu.srmuniversity.ac.in | HARSH GUPTA 81 | 2023-2-14 | Link | VERIFIED hgreatgupta2001@gmail.com | Harsh Gupta | 2023-2-8 | Link | UN-VERIFIED vinitjoshi.2345@gmail.com | Vinit Joshi | 2023-2-15 | Link | VERIFIED
5.	Anyone can check status by just 3- clicks on portal.	In a section called 'Verify Documents' on website, entities can upload any document and can see document status.	 Welcome to the Platfrom Upload the file to check the authenticity. Choose File BD-and-ML-Fundamentals-v3.png SUBMIT Results for Upload File Verified Status Secondary text: true Uploader Name Secondary text: Test j Uploader Email Secondary text: test002forest@gmail.com Upload Date YYYY-MM-DD: 2023-5-15

Chapter 5

CONCLUSIONS AND FUTURE WORK

5.1 FINDINGS

From project work completed till now, we observed that:

- All three categories of users of the platform can use the home page.
- Users can sign up and log in with the help of Google Authentication.
- Unique hash created for every unique document (in backend) and similar documents can be recognized using verify document section.
- Authorizer can preview the document and choose whether he/she should authorize it or not.
- Authenticity of any document is just a few clicks away from any user across the globe.

5.2 CONCLUSION

While working on our major project, we were successful in accomplishing the following objectives:

- To create profile metadata for admin/user to upload or verify the document- like owner, social link, email.
- To create digital fingerprint of documents by using hashing algorithm SHA-256.
- Users can put verification request of their documents to admin.
- Admin can preview verification documents to decide whether he should authorize it or not.
- Any entity from entire globe can check documents status just by 3-clicks on portal.

5.3 FUTURE WORK

For future work, we will be working on achieving the following objectives:

- To make User Interface more interactive and modern.
- To optimize the functionality of Platform.
- To develop it as a suitable alternative to DigiLocker (Indian Government App) on the basis of scalability and better domain coverage.

REFERENCES

[1] "University Grants Commission," [Online]. Available: https://www.ugc.gov.in/oldpdf/consolidated%20list%20of%20all%20universities.pdf. [Accessed 12 April 2023].

[2] The Hindu, "UGC issues list of 21 fake universities," The Hindu, 26 August 2022. [Online]. Available: https://www.thehindu.com/news/national/ugc-declares-21-universities-as-fake-maximum-in-delhi-followed-by-uttar-pradesh/article65814063.ece. [Accessed 12 April 2023].

[3] "Fake credentials in India: Problem, victims, and solution.," TruScholar, 04 April 2023. [Online]. Available: https://www.truscholar.io/fake-credentials-india-problem-victims-solution/. [Accessed 12 April 2023].

[4] M. J. Rodgers, "Education Verification for Employment: A Complete Guide [2023]," Iprospectcheck, April 2023, [Online]. Available: https://iprospectcheck.com/education-verification/. [Accessed 12 April 2023].

[5] H. Hockx-Yu, "Digital preservation in the context of institutional repositories.". [Online]. Available: http://eprints.rclis.org/8189/1/DPinIRs_Final.pdf. [Accessed 12 April 2023].

[6] C. Lakmal, S. Dangalla, C. Herath, C. Wickramarathna, G. Dias and S. Fernando, "IDStack — The common protocol for document verification built on digital signatures," 2017 National Information Technology Conference (NITC), 2017, pp. 96-99, doi: 10.1109/NITC.2017.8285654.

[7] Tetali, Rama & Reddy, P. & Rayudu, Srinivas & Raghavendran, Dr. Ch V & Lalitha, R. & Annapurna, B.. (2021). Proposing a reliable method of securing and verifying the credentials of graduates through blockchain. EURASIP Journal on Information Security. 2021. 10.1186/s13635-021-00122-5.

[8] A. Jamal, R. A. A. Helmi, A. S. N. Syahirah and M. -A. Fatima, "Blockchain-Based Identity Verification System," 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET), 2019, pp. 253-257, doi: 10.1109/ICSEngT.2019.8906403.

[9] Devdoot Maji, Ravi Singh Lamkoti, Hitesh Shetty, Bharati Gondhalekar, 2021, Certificate Verification using Blockchain and Generation of Transcript, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 03 (March 2021).

[10] Marella, Venkata and Vijayan, Anoop, "Document Verification using Blockchain for Trusted CV Information" (2020). AMCIS 2020 Proceedings. 12. https://aisel.aisnet.org/amcis2020/adv_info_systems_research/adv_info_systems_research/12.