[redacted]

A Final Year Report on

"Document Management System (PAPERBANK)"

By

Aayam Dahal
(T.U. Exam Roll No.15738/074)

An Internship report submitted in partial fulfillment of the requirement for the degree of Bachelor of Science (B.Sc.) Computer Science and Information Technology Awarded by IOST, Tribhuvan University

Sagarmatha College of Science and Technology

Sanepa, Lalitpur
September 2022

Sagarmatha College of Science and Technology
Sanepa, Lalitpur
(T.U. affiliate)

MENTOR'S RECOMMENDATION

I hereby recommend that this report has been prepared under my supervision by **Mr. Aayam Dahal**, in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology, be processed for evaluation.

.....

Admond Tamang

Senior Developer

General Technology Pvt. Ltd.

Naxal, Kathmandu

Sagarmatha College of Science and Technology
Sanepa, Lalitpur
(T.U. affiliate)

SUPERVISOR RECOMMENDATION

I hereby recommend that this report has been prepared under my supervision by **Mr. Aayam Dahal** in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology be processed for evaluation.

.....

Signature

Pratik Timalsena

Sagarmatha College of Science and Technology

Bsc.CSIT Department

Sanepa, Lalitpur

Sagarmatha College of Science and Technology
Sanepa, Lalitpur
(T.U. affiliate)

LETTER OF APPROVAL

The undersigned certify that they have read and recommended to the Department of Computer Science and Information Technology for acceptance, an internship report submitted by **Mr. Aayam Dahal** partial fulfillment for the degree of Bachelor of Science in Computer Science & Information Technology.

.....

Signature

Ramesh Singh Saud

Tribhuvan University

(External Examiner)

.....

Signature

Pratik Timalsena

Sagarmatha College of Science and Technology

(Project Supervisor)

.....

Signature

Er. Manish Aryal

Sagarmatha College of Science and Technology

(Program Coordinator)

ACKNOWLEDGEMENT

I am glad to take opportunity to express my sincere gratitude to supervisor of internship program Mr.Pratik Timalsena for his guidance, support and valuable suggestions and coordination until internship program was complete. I express my sincere gratitude to him for believing in me and giving me the chance to complete this internship and helping me in the time of uncertainty. It was indeed a great learning experience.

I am very thankful to Er. Manish Aryal, B.Sc. CSIT Coordinator of Sagarmatha College of Science and Technology, for providing us regular feedbacks and suggestions to perform an internship project assignment so that we can practically implement our studies in the real world.

I express my sincere gratitude to Mr.Admond Tamang for mentoring me throughout the entire development of project. Without his help we mightn't be able to convert our dream to working system.

I am very grateful to General Technology in providing all necessary help and support for entire development of project as well as infrastructures and facilities during our internship period. We are very thankful for their cooperation, without which it would never have been possible to complete this project.

Last but not the least, I am also thankful to all my teachers and friends for their valuable ideas, suggestions and support to complete this project.

Sincerely,

Aayam Dahal (15738/074)

ABSTRACT

The internship days spent at General Technology Pvt. Ltd located at Naxal, Kathmandu, provided a great chance of being connected more to the IT world and to learn how software is developed professionally. This report consists of details regarding the development of “Document Management System” which is a web based document management system still in development and used for the purpose of managing the documents and attachments within a financial institution like banks.. This system has a simple yet very rich user interface to help users get a better experience using this platform.

Keywords: *DMS, OCR, FTP, API.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HTML	Hypertext Markup Language
CSS	Cascading Style Sheet
JS	Javascript
SQL	Structured Query Language
DMS	Document Management System
DBMS	Database Management System

CHAPTER 1: INTRODUCTION

1.1. Introduction

"Document Management System" is a Tailored Software which is developed to manage, store and track electronic documents such as PDFs, Word processing files or paper-based information which are captured using document scanner. It is a software solution for organizing, securing, capturing, digitizing, tagging, and approving documents. The Document Management System will help vendors to enhance security by keeping sensitive data such custom information, their account information protected which is critical for any organization. It is feature packed with OCR, Encryption, Documenting Indexing, watermark, uploading attachements.

Internship period is regarded as a stepping stone in building one's career in their specific field. Interns not only get familiar with the programming languages but learn the practices that are carried out in a software company. This period lets one to learn working in a team and learn how the software development is carried out in the industry. More Importantly, it helps to get habituated to working environment before working in the industry. The theoretical part is slightly different when it comes to working. So, the internship period is most important. The opportunity to join General Technology Pvt. Ltd.. was gained and assignment to the frontend and backend team was provided.

1.2. Problem Statement

In today's world technology like a Document Management System is really important in the corporate world of banks. It is a hassle in today's world to organize documents in shelfs, racks, and buildings so to overcome this problem DMS that our company uses is feature packed with OCR, encryption, bulk-upload and many more.

1.3. Objectives

The main objectives of the Document Management System are:

- To implement maker and checker flow with OCR search and bulk upload documents into FTP server.
- To implement document scanning and indexing, enable maker checker notification, generate dynamic report and implement fuzzy and full text search.

1.4. Scope and limitation

Document Management System allows vendors to enhance security by keeping sensitive data such as customer information, their account information in fully protected which is critical for any organization. It also helps vendors to reduce storage space. Commercial property costs are increasing and so is the expense to store paper documents. A software-based DMS that can reduce the need for cabinets, boxes, and storage bins which is a valuable aspect to any enterprise, freeing up precious office space. It will also retrieve documents easily. DMS can retrieve files by words, phrases, tags, document types, location etc.

1.4.1. Scope of the Project

The scope of this project is as follows:

- This system makes it easier to manage documents and attachments.
- This system makes it easier for users to index documents, organize hierarchical documents and perform OCR search.

1.4.2. Limitation of the Project

- Our system is only web-based system and not android or any other mobile app based.

1.5. Report Organization

The internship report on "**Document Management System**" is based on four chapters. The overall report is organized into four chapters.

Chapter 1: Introduction

The first chapter provides an introduction to the internship project along with the problem statement, objectives, and scope and limitations of the project.

Chapter 2: Organization Details and Literature Review

The second chapters present the introduction to the organization, its organizational hierarchy along with its working domains and the description of the intern line. Similarly, it also includes review of similar projects and theories.

Chapter 3: Internship Activities

The third chapter consists of the roles and responsibilities acquired during the internship period along with the list of technical activities performed. It also provides the description of the project involved.

Chapter 4: Conclusion and Learning Outcomes

The final chapter is the conclusion section which includes the learning outcomes from the internship.

CHAPTER 2: ORGANIZATION DETAILS AND LITERATURE REVIEW

2.1. Introduction to Organization

General Technology Private Limited "GenTech" is founded in Kathmandu, Nepal in November 2011. GenTech is one of the highly regarded software development services and consultant, in the industry for its superior quality and commitment to innovation. GenTech inherently experienced in Information Technology Consultant, IS Audit, Document Management System (DMS), Business Process Automation (BPA), Business Process Outsourcing (BPO), Enterprise Document Management and Medical Solution.We help our clients augment their in-house development team, transform their ideas into working software products and connect their legacy enterprise systems to digital channels.

2.2 Organization Hierarchy

General Technology in hierarchy manner CEO works above two project managers which of them are classified as Web development and ATM / Hardware us as an interns work under project manager of web development along with other developers and QA.

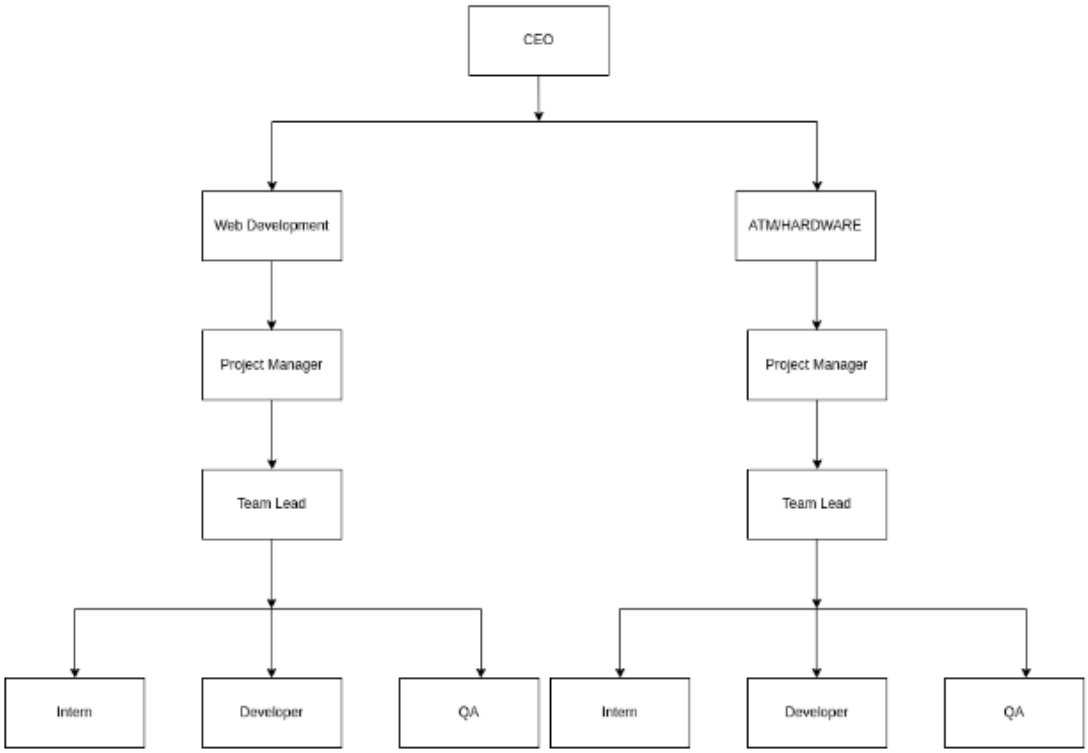


Figure 1: Organizational hierarchy

2.3. Working Domains of Organization

General Technology specializes in a broad range of Business Process Automation in Financial Institutions.. The following are the working domains of the organization:

1. Business Process Automation (BPA).
2. Document Management System (DMS).
3. Operation Risk Mangement System (ORMS).
4. Card Reconciliation System (ReCon).
5. Card Center Operation and Management System (CCOMS).
6. KIOSK Application.

2.4. Description of Intern Department/Unit

I worked in the web development and management department during my internship. GenTech has an extensive experience in building Web applications for organizations, both large and small. The products built are all scalable web applications with data processing back-ends (GenTech). Gentech's development methodology for Web applications follows these steps:

i. Gather requirements and Designers create several variations of static mockups.

ii. Designers make iterative changes to mockups based on client feedback to create a final version of mockup.

iii. Designers create clickable mockups to simulate a functioning product.

iv. Begin Web application development following a Model View Controller (Services) approach.

2.4.1. Duration of Internship

Table 1: Duration of Internship

Appointment Date	10th Falgun, 2078
End Date	10th Bhadra, 2079
Total Duration	6 months
Position	Intern as Full Stack Developer
Mentor	Mr.Admond Tamang
Office Hour	9:00 am – 6:00 pm

2.5. Literature Review/Related Study

Document Managment is an important aspect of any organization. Efficient document management allows the institution to be more organized. With the system,, employees could digitally access the documents, perform indexing and search the documents through fuzzy and elastic search. A Document Managements System is an integral part of any business as employees are the company's greatest asset. In fact, a well-structured and actionable Document Management System serves as support to organizing, indexing, uploading attachments, document encryption, digitization and many more. Some of the resources I used to accomplish the various tasks that were assigned to me are as follows:

2.5.1 React Hook Form

React Hook Form takes a slightly different approach than other form libraries in the React ecosystem by adopting the use of uncontrolled inputs using ref instead of depending on the state to control the inputs. This approach makes the forms more performant and reduces the number of re-renders. (Ail, 2022) In this project, react hook form was used to handle time consuming events such as onBlur, onChange, touch and onSubmit. All of these events are handled automatically by the react hook form.

2.5.2 Mailing service for notifications

Mailing is done with the use of nodemailer. Nodemailer is a module for Node.js applications to allow easy as cake email sending. The project got started back in 2010 when there was no sane option to send email messages, today it is the solution most Node.js users turn to by default.

2.5.3 Hashing password for authentication

The hashing of passwords is done with the bcrypt provider. bcrypt is a password-hashing function designed by Niels Provos and David Mazières, based on the Blowfish cipher and presented at USENIX in 1999. Besides incorporating a salt to protect against rainbow table attacks, bcrypt is an adaptive function: over time, the iteration count can be increased to make it slower, so it remains resistant to brute-force search attacks even with increasing computational power.

2.6 Data Modeling and Process Modeling

2.6.1. ER Diagram

ER-diagram represents how actually the database is designed. User, Admin, Documents and Attachments are major entities of the system. Several other entities sets are interrelated. Overall entity relationship is shown in the figure.

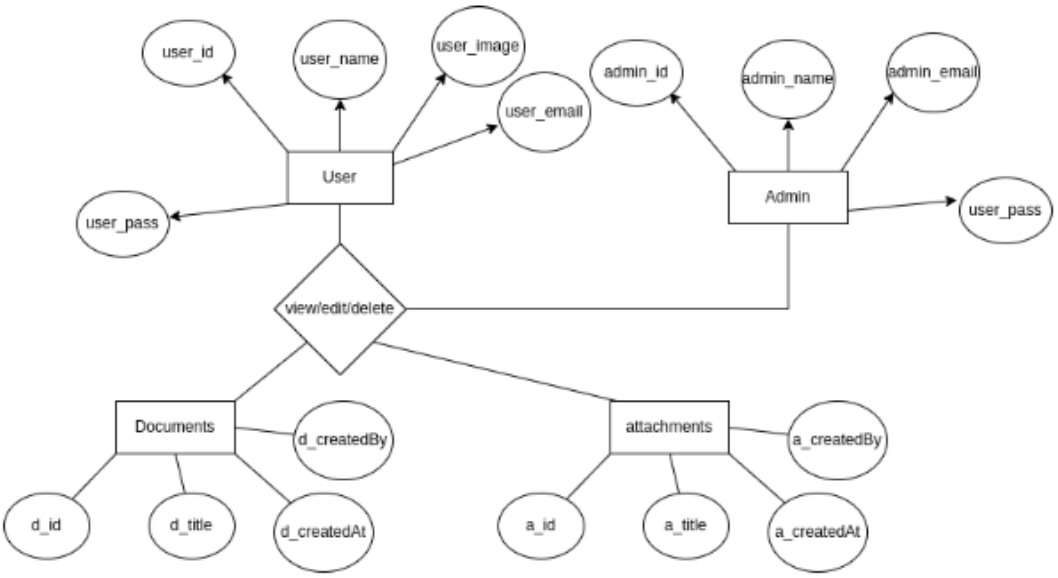


Figure : ER Diagram

2.6.1 Data Flow Diagram

Data Flow Diagrams show the flow of data from external entities into the system, and from one process to another within the system. DFD of my system is illustrated below

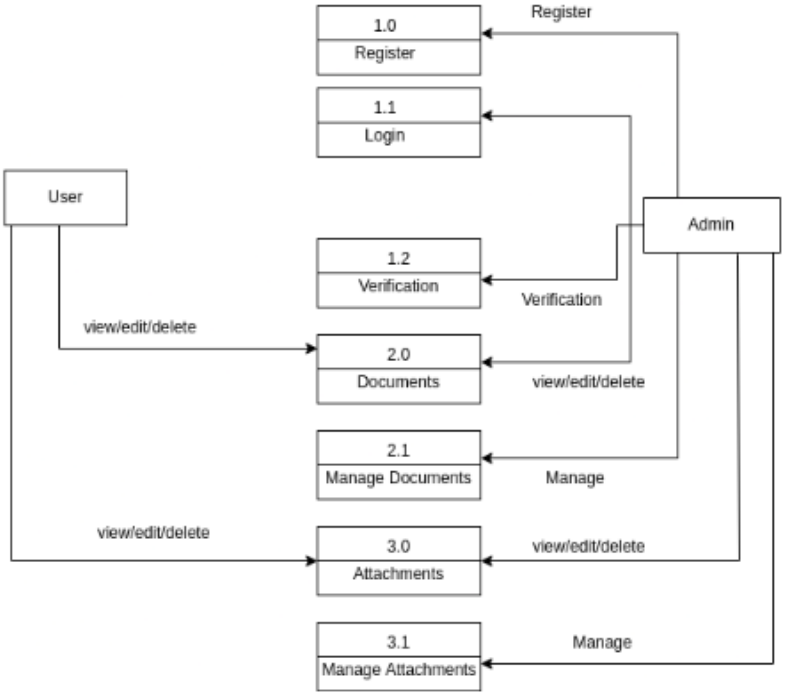


Figure: DFD Diagram

CHAPTER 3: INTERNSHIP ACTIVITIES

3.1. Roles and Responsibilities

I was assigned tasks by my mentor. Also, a weekly report is updated. All the tasks carried out in a week is updated on a regular basis on JIRA.

The responsibilities assigned to me by my mentor to work on Hierarchy and bulk upload module of "Document Management System" project where an admin can login, and create roles based on bank hierarchy such as operation, Maker, Checker, Branch Admin and so on. In this module, I had to create a hierarchy based permission for document edit/view/delete. On the bulk upload module multiple indexing of document had to be done so I worked on this module too. User could after index documents easily.

Some of the roles and responsibilities during my time were:

- DMS was a beginning project when I joined the company; I was first required to setup the project starter files on my device and understand the starter files and project structure.
- Research and implement the tasks assigned by my mentor/manager.
- Develop new features as per the design.
- Maintain the source code and its changes in GitHub.
- Fix the existing bugs and deal with any future bugs that may arise.
- Attend bi-weekly meetings to discuss the project progress, the tasks left to be carried out and participate in planning the next sprint activities.

3.2.Weekly Log

Table 2 : Details of the work done

Week	Activity
1st week	Project Setup and Overview of project. Install pnpm and python packages Tutorial on Formik, Tailwind and Yup validation (React Packages)
2nd week	Tutorial on LDAP, Windows Server and FTP. Setup LDAP client on vendor's server. Fix Hourly Access permission on vendor's server. Create a dynamic form using Formik and Tailwind.
3rd -6th week	MSSQL server setup. Database design overview. Create user group routes, and assign user on user group based on role verification. (User Group module allows the user of DMS to transfer documents to other user if an employee leaves the bank). Add Tailwind Components
7th -8th week	Citizen Bank visit for new deployment. Training on DMS to Citizen Bank employees. Started work on Bulk Upload Module. Fix OCR module for document with more than 3000 pages.
9th – 14th week	Add Indexing on Bulk Upload Added Tailwind components on React for logs table Added Reporting tools using Metabase UI changes Add Scanner priviliges on Roles section
15th – 26th week	Studied VAPT reports for DMS. Manage session and cookies. Remove sensisitve information from local storage. GIT traning classes

3.3. Description of the Project Involved During Internship

I was involved in developing the Document Management System application for General Technology during my internship. DMS helps to reduce the effort for managing physical documents by converting them digitally. The documents then can be indexed and encrypted in a FTP server further more it can be distributed in a hierarchical order which is alike to financial institutions like Banks.

1. **FTP file server**

All the documents and attachments and uploaded to a file server in the server. The documents and attachments that are uploaded are encrypted in the system.

2. **OCR search**

Optical Character Recognition is one of the features of the system. OCR (optical character recognition) is the use of technology to distinguish printed or handwritten text characters inside digital images of physical documents, such as a scanned paper document.It is implemented using Python and its packages. One of the packages used in this module is Tesseract.

3. **Hierarchy Based Permission & User Group**

Document Management System that I was working on was specific to banks. This module allows the vendor to define documents based on hierarchy of the banks. A simple example could the document uploaded on one branch cannot be seen on another branch's user. Another feature named user group is able to transfer document from one branch to another. The tools and technologies used in this project are:

- SASS
- Axios
- React Hook Form
- MSSQL
- Tailwind
- Redis
- pm2 for process management.

3.4 Tasks/Activities Performed

This report reflects 26 weeks of my internship on Web Application Development, and other technologies related to the subject. Before starting the internship, I was already interested in the topic and specially to have experience in a software development company. So, it was a great opportunity for me to gain knowledge and experience on these topics. As an intern, I got to work on lot of activities which are listed below:

- Setup project using command line.
- Create user group routes, and assign user on user group based on role verification. (User Group module allows the user of DMS to transfer documents to other user if an employee leaves the bank).
- Fix OCR module for document with more than 3000 pages.
- Studied VAPT reports for DMS.
- Integrated custom reporting tool, Metabase on RBB server.
- Fix watermark for individual banks like Citizen, BOK, RBB.
- Add Scanner privileges on Roles section.
- Added Tailwind components on React for logs table.
- Remove sensitive information from local storage.
- Implemented Passport Authentication.
- UI Changes according to vendor needs.
- Setup FTP server on vendor's server.
- Setup LDAP client on vendor's server.
- Integrate Tesseract for OCR module.
- Learned to use Git for version control, how to write proper git commit messages, how often git commits should be made, the process of creating branches and pushing into the branch.
- Used Redis server and MSSQL server to work in the local environment
- Used React Routing to redirect pages in single page applications, also implemented protected route.

3.5 Implementation

The “Document Management System” application is developed by constructing a customized User Interface by targeting to provide simple and user-friendly interface to the users. The user-friendly interface is maintained so that users can use this application without any difficulties.

3.5.1 Implementation Tools

- **React**

  React is a free and open-source front-end JavaScript library for building user interfaces based on UI components. It is maintained by Meta and a community of individual developers and companies.

- **Node.js**

  Node.js is a free, open-sourced, cross-platform JavaScript run-time environment that lets developers write command line tools and server-side scripts outside of a browser.

- **MSSQL**

  Microsoft SQL Server is a relational database management system developed by Microsoft. As a database server, it is a software product with the primary function of storing and retrieving data as requested by other software applications

3.5.2 . Implementation Modules

A. Document Index Module

In this module the admin first login into the system then he is authenticated user. If he is not valid then the message should be displayed saying that not valid. If he is valid then the page should be navigated to the “Admin Dashboard” displaying sections like “Documents, Archieved, Saved, Document Index, Security Hierarchy, and many more ”. If the user wants to create a document then it should navigate to “Documents” section and click on add document. Then the user will be navigated to a container

where s/he can input document details and check features such as OCR, Encryption and so on after this the user will be able to upload files to the FTP server and index the documents. Edit/Delete documents are controlled through roles and permissions.

B. Bulk Upload Module

In this module, the admin is able to upload more than two documents in a bulk and is able to index bundle of documents of same type. For e.g, An employee from a bank is required to upload Account Creation documents of customer s/he is able index the document based on Father Name, Mother Name and other details required during account creation. This module is efficient for uploading and indexing bundle of documents of same kind. After uploading relevant documents the user is able to enable features like OCR and Encryption.

3.6 Testing

Various types of testing procedures were performed in order to check the working mechanism and correctness of the system. Some of the types of testing that we did attempt are described below:

Test Case for Login

S.N	Test Cases	Input	Expected Output
1	Test if admin is able to login successfully.	Username: admin Password: admin123	Admin Dashboard
2	Test if unregistered user is not able to login to site.	Username: admin1 Password: admin123	Username admin1 is not registered
3	Test with valid username and empty password such that login must get failed.	Username: admin1 Password:	Password is required.
4	Test with empty username and empty password and check if login fails.	Username: Password:	Username and Password is required.
5	Test with empty username and valid password.	Username: Password:admin123	Username is required.
6	Verify the URL without loggind into the site.	----------	This page is not accessed.

Table 3: Test Case for Login

Test Result for Login

S.N	Input	Actual Output	Remarks
1	Username: admin Password: admin123	Admin Dashboard	Test Succeed
2	Username: Password:	Username and Password is required.	Test Succeed
3	Username: Password:admin123	Username is required.	Test Succeed
4	Username:admin Password:	Password is required	Test Succeed
5	Username:aayam Password:admin121	Username or Password incorrect	Test Succeed

Table 4: Test Result for Login

Test Case for Bulk Upload

S.N	Test Cases	Input	Expected Output
1	Test if admin is able to upload attachments.	Click on upload button and select files	Uploaded Successfully
2	Test if uploaded documents can be indexed	Select Document Index dropdown	Indexing options like: 1. Account Number 2. Account Type 3. Citizenship 4. Branch Code
3	Test if attachments are indexed	Select Document List	Indexed Documents.
4	Test if documents have relevant watermark	Click on Select Files	Documents should have relevant watermark of the vendor.

Table 5: Test Case for Bulk Upload

Test Result for Bulk Upload

S.N	Input	Actual Output	Remarks
1	Click on upload button and select files	Uploaded Successfully	Test Succeed
2	Select Document Index dropdown	Indexing options like: 1. Account Number 2. Account Type 3. Citizenship 4. Branch Code	Test Succeed
3	Select Document List	Indexed Documents.	Test Succeed
4	Test if documents have relevant watermark	Documents should have relevant watermark of the vendor.	Test Succeed

Table 6: Test Result for Bulk Upload

CHAPTER 4: CONCLUSION AND LEARNING OUTCOMES

4.1 Conclusion

As a full stack developer, I was able to learn a lot about the real world execution of different projects. There were a variety of tasks assigned to me which ranged from creating reusable components, writing optimized code, unit testing and working in the UAT, creating responsive web pages, api development and api testing. Throughout the internship period, I have tried my best to learn from a professional aspect and implement what I have learned in the BSc.CSIT course. The internship course helped me to explore the different career paths which appealed to me. Moreover with this experience I learned how to function within a professional workspace.

4.2 Learning Outcomes

My internship period at General Technology turned out to be immensely fruitful. Through this internship, I got familiar with the working environment of an IT company, new technologies, working under pressure and to work with a team. Concurrently working on multiple features and handling bugs & exceptions within the given deadline has raised my confidence and also enhanced my software development skills. Besides this, the extent of my expertise and solving problems through practical reasoning has improved to a heightened level. Apart from this, I was made aware on how developmental strategies for projects are fabricated, how work monitoring is carried out and human resource management is carried out in an IT company. Above all, I learned the importance of time management, team management and persistence to solve problems.

References

React Documentation (2022, January). *React Documentation*. Retrieved from reactjs.org: https://reactjs.org/docs/getting-started.html

Mdn Web Docs. (2022,February). *JavaScript Documentation*.Retrieved from MDN docs: https://developer.mozilla.org/en-US/docs/Web/JavaScript

Python. (2022, January). *Python Documentation.* Retrieved from python.org : https://www.python.org

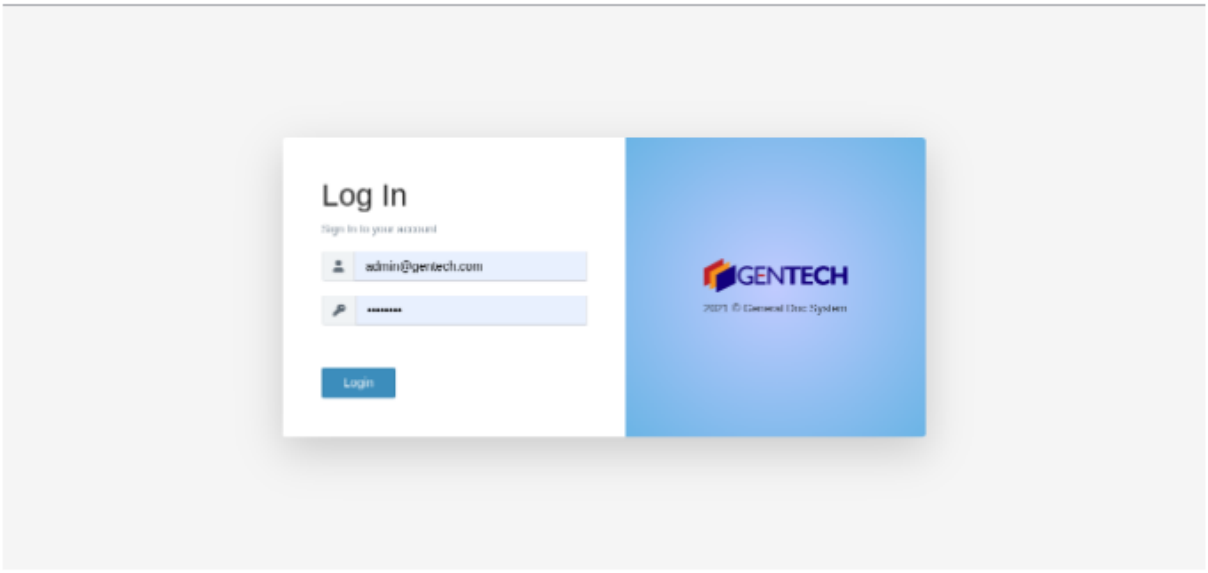
Stack overflow. (2022, January). Retrieved from Stack Overflow: https://stackoverflow.com/

Tailwind UI. (2022, January). *Tailwind Documentation.* Retrieved from Tailwind UI: https://tailwindui.com/

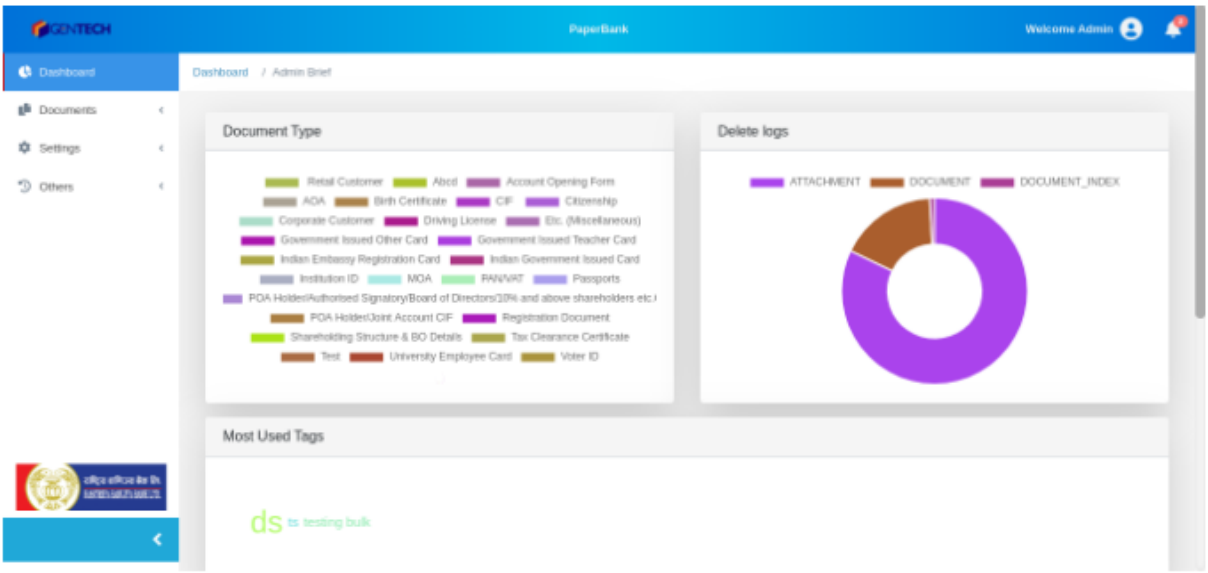
Appendix

A. Snap Shots

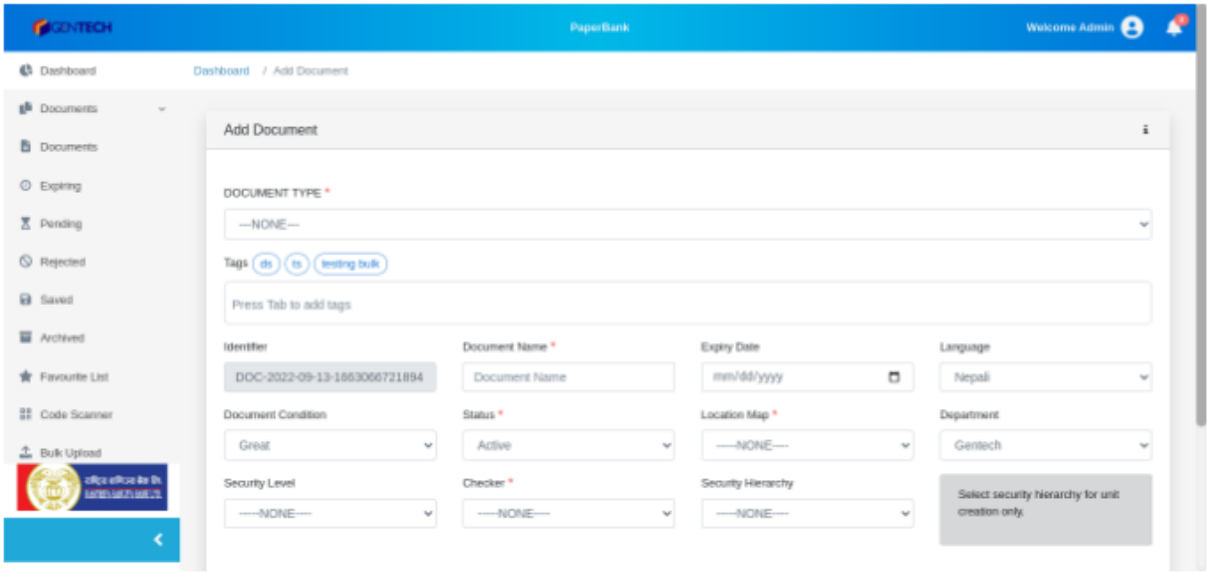
Login Page



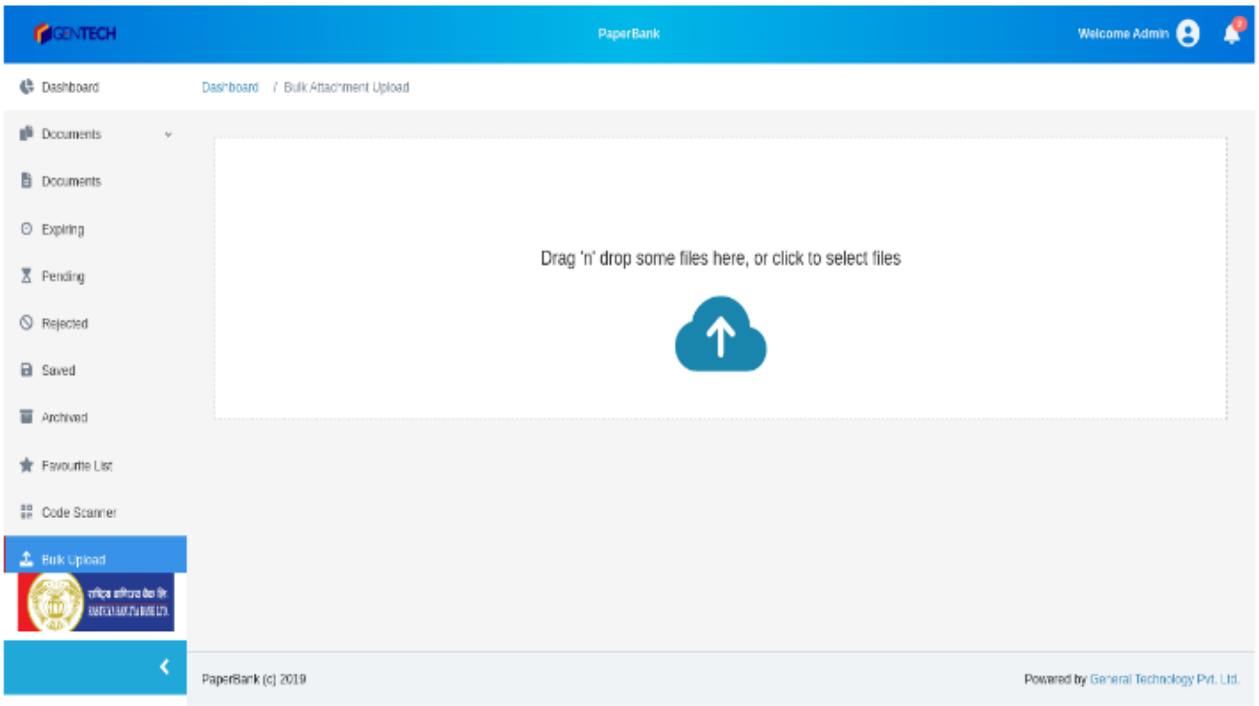
Admin Dashboard



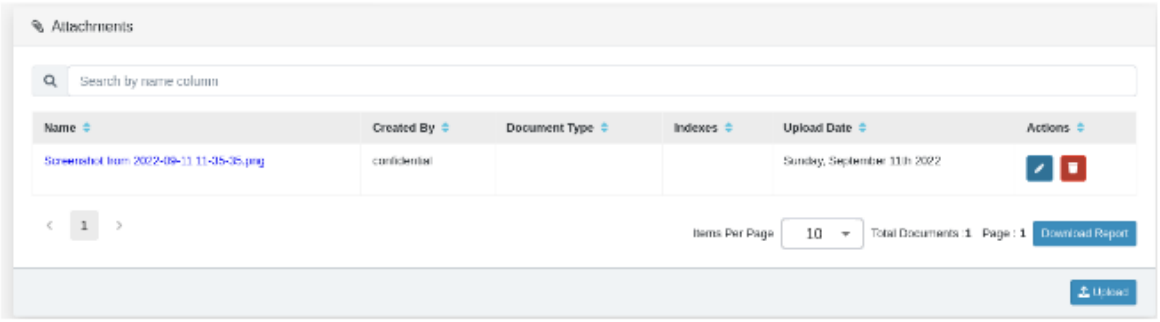
Add Document Section



Bulk Upload Section



Attachment Upload Section



Attachment View Section