**DESIGN AND IMPLEMENTATION OF A WEB BASED PROJECT REPOSITORY WITH PLAGIARISM CHECKER**

**BY**

**TUNDE-IDOWU INIOLUWA**

**ADEYEMI BANJI**

**EMORDI DIVINE**

**A PROJECT WORK SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF SCIENCE B.SC(HONS) DEGREE IN SOFTWARE ENGINEERING**

**TO THE DEPARTMENT OF SOFTWARE ENGINEERING**

**SCHOOL OF COMPUTING AND ENGINEERING SCIENCE**

**BABCOCK UNIVERSITY ILISHAN-REMO**

**OGUN STATE, NIGERIA**

[**CHAPTER ONE**](#_w3abupvfab8w) **4**

[**INTRODUCTION**](#_9tsvicf72t4w) **4**

[1.0       BACKGROUND TO THE STUDY](#_9g3x60vjr18g) 4

[1.1       PROBLEM STATEMENT](#_on2nkh9397vr) 7

[1.2       AIM AND OBJECTIVES](#_9ktoxetfyjmt) 7

[1.3  METHODOLOGY](#_pebm69xkd5y2) 8

[1.4 SCOPE OF THE STUDY](#_30maomnlvjf) 8

[1.5   SIGNIFICANCE OF THE STUDY](#_oducp7fskimn) 9

[**CHAPTER TWO**](#_iccepewbkoki) **10**

[**LITERATURE REVIEW**](#_eovbh9zfem0h) **10**

[2.0 INTRODUCTION](#_flhbhjg1j8l7) 10

[2.1 CONCEPTUAL FRAMEWORK](#_l04a50wsqi6k) 11

[2.1.1 Digital Repository](#_1xrumz3nxxbu) 11

[2.1.2 Institutional repository](#_ypgt3kpna6kx) 11

[2.1.3 Disciplinary repository](#_irag9xn1p2wg) 11

[2.1.4 Plagiarism](#_hhxytvz5ymsa) 12

[2.1.5 Brief History of Digital Preservation with Repositories](#_twaw37gvobhv) 13

[2.1.6 Need for Plagiarism Checker](#_3kliorf36w3t) 13

[2.2 Review of Related Works](#_pv323wmkklwi) 14

[2.3 Gap Analysis](#_wetda2tj7so4) 18

[**CHAPTER THREE**](#_b6ha17v5k58l) **19**

[**SYSTEM ANALYSIS AND DESIGN**](#_24wic2c4f1a8) **19**

[3.0 INTRODUCTION](#_wd03v6cvdky1) 19

[3.1 ANALYSIS OF THE CURRENT SYSTEM](#_nh8l44bpvf7g) 19

[Figure 3.1 Analysis of the current System](#_gkblpw62o7ou) 20

[3.2 DESIGN AND IMPLEMENTATION OF A WEB BASED PROJECT REPOSITORY WITH PLAGIARISM CHECKER.](#_52uckak1sclk) 20

[3.2.1 INPUT DESIGN](#_cwo1evamoy3n) 20

[3.2.1.1 Signup Form](#_ije4vxexul65) 21

[3.2.1.2 Login Form](#_9t757tvky3sv) 21

[3.2.1.3 Update User Details](#_7f36m7d0x9td) 21

[3.2.1.4 Project Registration](#_rmj370pj3183) 22

[3.2.1.5 Search Bar](#_vt7ps0r32hit) 22

[3.2.1.6 Plagiarism Check](#_ui5w2c9b36a2) 23

[3.2.2 OUTPUT DESIGN](#_lj6cflmsh59z) 23

[3.2.2.1 Project Feed](#_cqe8w8e9pv7k) 23

[3.2.2.2 Project Details](#_mfjc5aryz222) 23

[3.2.2.3 Search Result](#_ynul7tngtujg) 24

[3.2.2.4 Plagiarism Report](#_q8ensli20gsw) 24

[3.3 SOFTWARE DEVELOPMENT PROCESS MODEL](#_7mpt41agyx9g) 26

[3.3.1 Phases of Iterative Development Process Model](#_vvhb9ld15zf9) 27

[Figure 3.3.1 Phases of Iterative Development Process Model](#_4ksrjmx1o3p6) 27

[3.4 REQUIREMENTS ANALYSIS](#_38ck4rlbaij5) 28

[3.4.1 Functional Requirements](#_fiu7svkdt8qm) 29

[3.4.1.1 User Requirements](#_v3ydgctfie7y) 29

[3.4.2 NON-FUNCTIONAL REQUIREMENTS](#_3lu9ewqcadly) 30

[3.5 SYSTEM DESIGN](#_skcpo6flpbqp) 31

[3.5.1 USE CASE DESIGN](#_5ck6ldxoupmw) 31

[3.5.2 DATABASE DESIGN](#_hiciruy4y2d5) 32

[3.5.3 MODULE DESIGN](#_udx3v13eu6n0) 33

[3.6 DESIGN AND DEVELOPMENT TOOLS](#_cny09s4jpzmq) 33

[**REFERENCES**](#_jn9fdwr2kdpg) **38**

# **CHAPTER ONE**

# **INTRODUCTION**

### 1.0       BACKGROUND TO THE STUDY

A digital repository is a collection of online resources. Preservation of research outputs and scholarly resources is required in order to do research in a particular sector of expertise.

The development of digital repositories as a crucial component of research infrastructure is accelerating.

Having a readily available online community of publicly accessible digital repositories provides a significant advantage for novice researchers to begin browsing and extracting resource materials/ideas to easily plan his/her research project. When there is a system that allows other research output to be showcased, it provides a significant advantage that can eventually converge into a single large online research community. The OAIS model offers a conceptual framework for creating repositories that adhere to standards. The early 2000s saw the emergence of the relatively new phenomena known as digital repositories. To make sure that all kinds of repositories receive top-notch preservation services, the idea of trusted digital repositories has been advanced. (Iris Xie, Krystyna K, 2016).

Most institutions have little to no access to an online digital repository of their research work, which can be shared with other researchers for collaboration. Other online digital repositories are either privately accessible to institutional stakeholders or require a paid subscription. Furthermore, approximately 80-85% of university digital intellectual output is never made public (P. Jain, G. Bentley, and M. Oladiran, 2013).

Specific needs led to the division of repositories into disciplinary and institutional repositories, which contributed to the growth of digital repositories. A disciplinary repository is an online library with academic content for research projects in a certain field of study. A university provides its students, faculty, and staff with institutional repositories for the management and sharing of digital assets produced by the institution and its community. A repository that is more broadly referred to as a "open access repository" is one that makes the full text of items (or complete files) it contains freely and immediately available without any access limitations (Pinfield, 2005).

Pinfields' study, "The Role of Open Access Institutional Repositories," makes the case that research articles should be required to be submitted to open access institutional repositories since they enhance content dissemination. A global research archive's search services might be developed using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) technology, according to the study. Open access mandates are argued to be the best but not the only option to enhance academic communication, despite the advantages of open access repositories being highlighted, such as the increased impact potential for research publications. The Directory of Open Access Repositories -OpenDOAR maintained by the University of Nottingham is intended to help repository administrators in providing a better service for their users and facilitating repository growth. It provides a listing of open access repositories to enhance and support the academic and research activities of the global community (Basavaraja, 2017).

There is a shared agreement that plagiarism occurs when someone else's ideas, thoughts, or work is misrepresented as their own, whether intentionally or unintentionally, without appropriate acknowledgment. Plagiarism detection by hand can take a very long time. We require a software program that can identify plagiarism in order to lessen the workload placed on university lecturers. The challenge with identifying plagiarism is the sheer volume of content we are dealing with. Any system's processing capacity may not be able to handle the task of comparing a document with hundreds of other registered papers to identify plagiarism. When we have a large amount of data at our disposal, it can take months to find plagiarism. What difference, then, can there be?, the introduction of a system that can quickly and accurately identify plagiarism.

This study aims to introduce a system where research materials can be made available in a central resource center in form of a web-based digital project repository to store and retrieve research materials to solve the issue of availability and accessibility alongside a plagiarism checker intended to help reduce plagiarism and improve author’s integrity in written research work.

### 1.1   PROBLEM STATEMENT

The process of shelving and archiving past project or research works limits the level of accessibility  and consequently makes it almost impossible for students to leverage on the existing work for the enhancement of their research. . In view of this, developing a digital document repository for project or research works enhanced with plagiarism check becomes necessary.

### 1.2 AIM AND OBJECTIVES

 The project is aimed at designing and implementing a web-based project repository system with a plagiarism checker, whose objectives include:

  i.           Determine what constitute the best document repository and plagiarism checker method

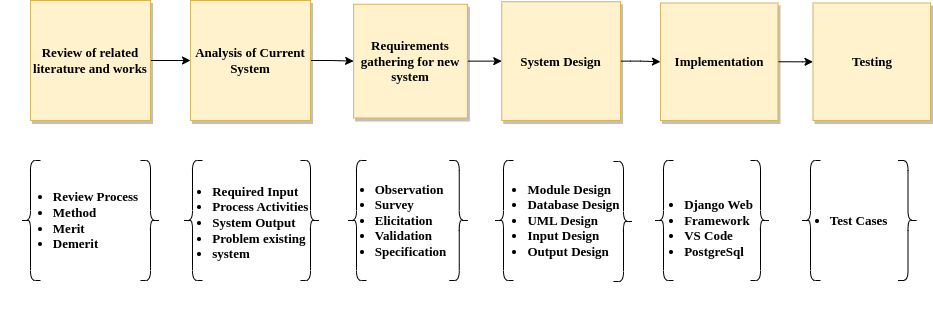
 ii.           Design the document repository and plagiarism checker system

  iii.          Implement and evaluate the document repository and plagiarism system

### 

### 1.3  METHODOLOGY

In our bid to achieve the above stated aim and objectives, the methods represented below were incorporated into the project’s life cycle.



### 1.4 SCOPE OF THE STUDY

The scope of this research project covers the design and implementation of a web based application for a project repository and integration with a plagiarism checker. The implemented system will be able to run on any browser over the internet.

Although this technology can be used all around the world and in any field of study, Babcock University, Ilishan-Remo would be used as a test case for the implementation of the model. The developed system will be able to collect and store scholarly materials as well as detecting plagiarism content on documents before upload

### 1.5   SIGNIFICANCE OF THE STUDY

One foremost significance of this research is the provision of a unique identifier to all the documents in the repository. Therefore making it easy with the aid of the search module incorporation into the repository application to efficiently search, discover and retrieve the required project from its database. Users will benefit from this feature in a much greater way by being able to search for a project either by the  title, author or year. The project repository could also serve as an educational environment for many internet users and researchers. It will encourage researchers to take up further research on a subject matter (like a project) and thereby add to the body of knowledge. This repository also covers authors of the projects and because it clearly attributes the student name (author) to its content.

### 

### 

### 

### 

### 1.6   ORGANIZATION OF PROJECT WORK

Chapter 1: This chapter has successfully examined the conceptual introduction of the design and implementation of a Web Based Project Repository with Plagiarism Checker. This chapter also identified the need for carrying out this study and highlighted the major objectives of the study under review, as well as the proposed methodology.

Chapter 2: This chapter briefly discusses major concepts and theories associated with the implementation of a Web Based Project Repository with Plagiarism Checker as well as the approaches adopted in related and existing systems alongside examples of similar systems already in place. It also discusses limitations of the stated systems.

Chapter 3: This chapter discusses the system analysis and design as they are applied to this project. This chapter analyzes the current system and the proposed system as well as methods of implementation.

Chapter 4: This chapter discusses the processes used to implement the system design, software solution and the management of other development constraints.

Chapter 5: This chapter is the last and acts as the conclusion. It includes a summary of the entire project, results, findings and recommendations for further study.

# CHAPTER TWO

# LITERATURE REVIEW

### 2.0 INTRODUCTION

The Web-based digital repository and plagiarism checker is a system set to aid convenience and ease accessibility to scholarly materials from various fields of study alongside the functionality of a plagiarism checker to detect plagiarism and improve author integrity. This provides an automated approach to research management and plagiarism detection rather than the manual methods of shelving and manual comparisons which aren’t efficient or effective.

Digital repositories are now more crucial for the collecting and dissemination of academic materials because of the growing trend toward online scholarly communication and the absence of scholarly content management systems among universities (Budapest, 2002; Chan, 2004; Lynch, 2005). At academic institutions today, digital repositories are employed to archive and disseminate university scholarly products (Lynch and Lippincott, 2005). The use of digital repositories give rise to a reoccurring educational challenge, that is, plagiarism detection and prevention. It is advisable for research institutions and individuals to utilize the use of a plagiarism detection system to help eliminate the practice (Patil, 2019).

This project is aimed at building a system which collects user details and scholarly materials such as articles, journals and stores them on a central server on the internet for ease of access and portability. It also provides the functionality of a plagiarism checker to detect theft of intellectual property as being able to freely access materials digitally also paves way for digital theft. The use of a plagiarism checker helps curb this.

This chapter discusses all major concepts behind the system and explains various theories that support this research work, also, reviewing approaches taken by previous researchers while developing a digital repository and plagiarism checker.

### 2.1 CONCEPTUAL FRAMEWORK

This section discusses the specific concepts that make up this research project.

#### 2.1.1 Digital Repository

Digital repositories are information systems that ingest, store, manage, preserve, and provide access to digital content.(Xie et al., 2016). A repository can perform a variety of tasks. It can be a location where particular databases, files, or documents are stored for access or dissemination. A repository can aggregate the data into a searchable storage location or it can enable the selective extraction of data (Vangie, 2022).

#### 2.1.2 Institutional repository

This is an archive for gathering, storing, and sharing digital versions of an organisation's intellectual output, typically by a research institute. By conserving and facilitating access to digital content, institutional repositories carry out the essential duties of digital libraries. They make it possible for scholars to archive their work and increase the impact and use of their research. Institutional repositories, as opposed to subject or theme-focused digital libraries and discipline-specific repositories, record original research and other intellectual output from an institution's constituent population that is active across a variety of fields. (Crow, 2002)

#### 2.1.3 Disciplinary repository

Disciplines differ significantly from one another in terms of topics, publication rates, information-seeking habits, and social norms. A disciplinary repository is an online archive that houses the writings of academics in a particular field of study or the data related to those writings. Scholars from any university may submit their work to disciplinary repositories. A disciplinary repository is similar to other repositories in that it collects, disseminates, and archives material, but it is more narrowly focused on a particular field. These collections can include academic and research papers. A disciplinary repository can offer specialised data, services, and tools used and preferred by a particular scientific community, in contrast to institutional and general repositories that serve a wide range of users and data with metadata, access, and user support mechanisms tailored to a heterogeneous and wide audience (Lyle, 2017).

#### 2.1.4 Plagiarism

The word “plagiarism” was first described in English in the year 1601 by the dramatist Ben Jonson, to describe someone guilty of literary theft (Valpy, 1999). Plagiarism is one of a number of research misconducts, which also include the fabrication of results, falsification of data, incorrect interpretation of data, making specific inferences, and the use of copied data or ideas inside a study report. A breach of confidentiality or a violation of authorship are both examples of misconduct. Research has developed into a booming sector in recent years. More than 7.1 million researchers worldwide compete fiercely to have their work published in more than 25,000 journals. There is pressure on researchers to publish their work in reputable publications. The prevalence of plagiarism rises when this pressure is combined with a lack of time, a lack of research abilities, and the simplicity of finding information and publications online (Rahab et. al 2015). Recent reports of plagiarism may be on the rise for one or more of the reasons listed below: The ease of information, the pressure to publish (publish or perish) in academia for career advancement, the lack of writing confidence and skill, especially among novices, the writing of manuscripts under pressure to meet deadlines, the ignorance of what constitutes plagiarism, and the ignorance of authors are all factors (Jyotindu, 2016).

#### 2.1.5 Brief History of Digital Preservation with Repositories

The early 2000s saw the emergence of the relatively new phenomena known as digital repositories. To make sure that all kinds of repositories receive top-notch preservation services, the idea of trusted digital repositories has been advanced. Major steps toward a thorough and team-based strategy for digital preservation weren't really made until the Internet and personal computer use exploded in the 1990s, which led to an exponential rise in the amount and variety of digital items ( Xie et. al, 2016).

#### 2.1.6 Need for Plagiarism Checker

The most prevalent issue with research writing is plagiarism. Increasing awareness of this issue is the key to stopping it from getting worse. Two crucial techniques for combating plagiarism in academic settings are holding workshops and offering plagiarism-detection software. The foundation of scientific research ethics is trust and honesty; plagiarism undermines these principles and fosters a climate of suspicion that impedes the advancement of science.

Major plagiarism that is found in a published article after it has been printed can have damaging repercussions for the reputations of the author, coauthors, journal editors and reviewers, as well as the institution to which the author belongs. The editor can impose a variety of penalties on the plagiarist author. These can include formal letters to the heads of the plagiarist's institution, refusal of future submissions from the author or his institution, retraction, and reporting. In cases of minor plagiarism that reflects a concept misunderstood, these can include a letter of explanation (Rahab et. al, 2015). (Shankar & Ramasesh, 2014) suggested that the main aim should be to enhance the originality of scholarly output.

### 2.2 Review of Related Works

(Smith et al., 2003) takes the Massachusetts Institute of Technology as a case study and emphasises the increasing need for a stable place in which a material may be stored and accessed long-term. Global statistics show that DSpace is the most widely used open source repository software for institutional and open access repositories. High use of the software has been observed in universities and research‐based institutions as a way to provide access to research output, scholarly publications, and more. Hewlett-Packard and MIT Libraries collaborated over two years to create the DSpace digital repository platform, an open source repository development software typically used for creating open access repositories for scholarly and published digital content. The organisation of objects on the DSpace repository into communities and collections- At an institutional level, communities could be departments, labs, research centres and collections in turn could contain related items accessed by users as scholarly materials in form of research articles, theses, dissertations or technical reports, provides the repository with an easy to navigate structure. The DSpace Repository is an open source system implemented using the Java programming language and PostgreSql as a relational database.

(Adegbile, et. al, 2021) emphasised the application of Project Repository System to educational administration as an alternative to the manual method of storing past project documents and class materials. Their study focused on the institution's past project works with the goal of minimising the stress, mistakes, loss, and other harms brought on by the manual technique of maintaining past project works in educational institutions. (Adegbile, et. al, 2021) designed and developed a web-based digital repository with an embedded search engine which was implemented using Visual Basic and Windows form application for the Federal College of Animal Health and Production Technology, Moor Plantation Ibadan, Nigeria.

(O’Brien, 2019) describes researchgate as a for-profit, social media–like scientific networking and collaboration website. By registering on researchgate, researchers are able to build profiles. The typical components of a profile include an overview similar to a dashboard, references to published work, contact and professional information, research interests and links to references of potential interest. A question submission-and-response knowledgebase, which enables members to post, react to, and track questions about research and other areas of interest, is one noteworthy feature of this system. (Lex, 2015) presents an assessment of the ResearchGate score as a measure of a researcher’s scientific reputation and concludes that the ResearchGate Score should not be considered in the evaluation of academics in its current form. The score is intransparent and irreproducible and changes made cannot be reconstructed despite its service as a measure of scientific reputation of a researcher.

# 

(Del, 2020) designed and developed a plagiarism detector that can register documents, provide people access, and determine how similar two documents are. Thus, HTML, PHP, JavaScript, CSS, and MySQL were used to create the software. Three key modules make up the developed system: Document Search, which lets users view documents, Document Registration, which lets administrators add and manage the stored papers, and Document Comparison, which acts as the system's plagiarism detection mechanism. Boyer-Moore Algorithm was employed to draw attention to the allegedly plagiarised document while Normalised Compression Distance was utilised to gauge similarity.

Richard Price established the social networking site Academia.edu in 2008. Over 52 million users have accounts with the service, which receives 36 million monthly unique visits. The site is less well-known in the social sciences and economics than ResearchGate is, and it is more well-known in the arts and humanities. In terms of the socioeconomic stratum, San Francisco is the corporate headquarters of Academia.edu, a for-profit organisation. Users grant Academia.edu a worldwide, revocable, non-exclusive, transferable licence to exercise any and all copyrights when posting, uploading, publishing, submitting, or transferring member content.

The platform's revenue model is focused in large part on offering a variety of free features that are complemented by premium accounts, like better analytics and a job board for posting academic openings. The homepage offers a continuous news feed that informs visitors of new uploads, bookmarked publications, and various user actions on a technological and cultural level.

Additionally, the homepage offers tools like Suggested Sessions and Suggested Academics for fostering connections based on shared research interests. Additionally, Sessions, a feature of Academia.edu, enables users to set up an unique page where other academics can provide general comments on papers or line-specific notes. The website promotes the exchange of a variety of scientific products, such as articles, books, book chapters, as well as conference presentations and teaching material. (Stephanie, 2018)

(Chinedu et. al, 2020) describes plagiarism as an improper act in which someone uses another person's words and ideas without giving the author full credit. The majority of educational institutions in Nigeria currently store their students' project work on shelves, making it difficult to compare one or more projects to ascertain whether plagiarism has occurred. When done manually, selecting files one at a time and comparing them to other files to check for plagiarism is inefficient and time-consuming for everyone. (Chinedu et. al, 2020) developed a system which when put through a plagiarism test, finds 40% of the words in the thesis body to match. The JAVA programming language was used in the implementation. The JAVA development kit 1.7, Netbeans source code editor 7.4, MySQL database management 5.0 for back end support, and fireworks graphics editor for the visuals were utilised to complete this system.

Skyline Inc. produced independent anti-plagiarism software that can identify plagiarised text. It is a stand-alone desktop application for Microsoft Windows created with Visual C#.Net. It utilises the precise substring detection technique and is utilised in academic settings. It runs exclusively on Windows operating systems, which is one of its flaws, and it generates a lot of false positives. (Chinedu et. al, 2020)

(Shahabi, 2012) Sherlock is used to detect copyright infringement in essays, computer source code files, and other types of digital texts. Sherlock measures how similar two papers are by transforming the text it receives into digital signatures. A digital signature is a number created by converting a number of input words (three by default) into a string of bits and then connecting those bits to create a number. Sherlock is written in the C programming language, and before it can be installed on Unix/Linux or Windows, it needs to be compiled. Since it is a command-line program, it lacks a GUI.

### 2.3 Gap Analysis

In order to gain recognition as trusted, a repository has to have certain attributes that ensure the reliability and authenticity of stored information. Combining the studies on digital repositories andplagiarism, the above reviewed repositories do not cater for plagiarism, hence, there’s no integrity in the archive of materials uploaded on their repository. This project aims to bridge that gap by enforcing the use of a plagiarism checker before acceptance of materials on the repository.

# CHAPTER THREE

# SYSTEM ANALYSIS AND DESIGN

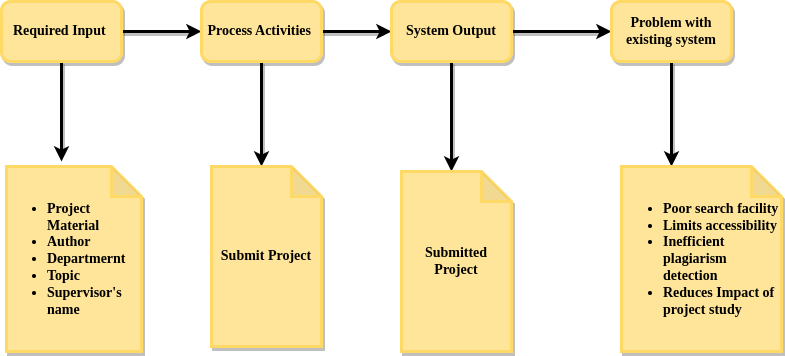
## **3.0 INTRODUCTION**

This chapter aims to provide a thorough explanation of the approach to be used in order to accomplish the aim and objectives of the proposed system. The proposed plans, process model, technologies, requirements, and system design that will be covered in this chapter are all included in the approach.

## **3.1 ANALYSIS OF THE CURRENT SYSTEM**

Taking a University institution as a case study, undergraduates and postgraduates are required to carry out a project study on a designated topic within a field of study. Students and researchers prepare a project document or journal article, printed in hard copy and submitted to the institution. Submitted printed scholarly materials usually have limited storage set out for them in the institution and much thought isn’t put in place to further manage these materials. The current system adopts storage in a physical location, hence, limits the accessibility of the materials from any location and the impact of the study for further research. Data related to these materials such as the authors, year of study, supervisor and topic are usually recorded on paper rendering the records vulnerable to physical damage and limits search time for easy accessibility.

Plagiarism in scholarly materials in a reoccurring educational challenge. The current system that caters for plagiarism detection is manual, hence, a time consuming and inefficient process that involves individual comparison of hard copy materials by hand.



##### Figure 3.1 Analysis of the current System

## **3.2 DESIGN AND IMPLEMENTATION OF A WEB BASED PROJECT REPOSITORY WITH PLAGIARISM CHECKER.**

The proposed system provides an adequate solution to ensure the integrity of scholarly materials using a plagiarism checker while also providing an easily accessible web-based repository to aid long-term preservation of these materials and impact on research. This system is solely a web-based application that allows user registration and login and carry out certain activities pertaining to project registration and plagiarism detection.

### 

### 3.2.1 INPUT DESIGN

This section contains the description of the various fields required as input.

#### 3.2.1.1 Signup Form

The user has the ability to register and create an account. The input fields required for this are as follows:

1. **First Name:** This field captures the first name of the user, for unique identification.
2. **Last Name:** This field captures the last name of the user, for unique identification.
3. **Email Address:** This field captures the email address for validation and notifications.
4. **Institution:** This text field captures the name of institution/organization which the user belongs to.
5. **Password:** This field captures the user input as a password for authentication.
6. **Confirm Password:** This field captures the user input as a password and ensures it matches with the password field.

#### 3.2.1.2 Login Form

This form is necessary to allow a user to gain access to the application by entering their email address and password.

1. **Email Address:** This is an email field containing the user’s email for identification.
2. **Password:** This is a password field that captures the user’s password for authentication.

#### 3.2.1.3 Update User Details

1. **First Name:** This field captures the first name of the user, for unique identification.
2. **Last Name:** This field captures the last name of the user, for unique identification.
3. **Email Address:** This field captures the email address for validation and notifications.
4. **Password:** This field captures the user input as a password for authentication.
5. **Confirm Password:** This field captures the user input as a password and ensures it matches with the password field.

#### 3.2.1.4 Project Registration

This form captures all relevant information pertaining to a project material.

1. **Project Title:** This is a text field containing the project’s title for which a research was carried out.
2. **Abstract:** This is a text area containing a brief summary of the project study… **DESCRIPTION**
3. **Department:** This is a drop-down field with specified departments of study under which a project was carried out.
4. **Tags:** This text field contains relevant keywords associated with the material.
5. **Year:** This is a number field that captures the year the project study was carried out.
6. **Author’s Name:** This field captures the name of the author of the material.
7. **Supervisor’s Name:** This field captures the name of the supervisor under which the study was carried out.
8. **Citation:** This section captures the citation information required for referencing the material.
9. **Document:** This is a file upload field that accepts only PDF documents pertaining to the project’s material.
10. **URL:** This is a URL field that captures any links relevant to the project.

#### 3.2.1.5 Search Bar

This is a single form field that captures information required to execute search queries such as

(i) Author’s Name

(ii) Supervisor’s Name

(iii) Department

(iv) Year

(v) Tag

#### 

#### 3.2.1.6 Plagiarism Check

1. **Document Upload:** this field allows the user to upload the document he/she wants to carry out a plagiarism check on.
2. **Supervisor’s Email Address:** This email field captures the supervisor’s email address.

### 3.2.2 OUTPUT DESIGN

This section contains the description of the various fields that would be delivered as output.

#### 3.2.2.1 Project Feed

This is a web page that contains all available projects in the repository. It renders projects in the repository once loaded.

#### 3.2.2.2 Project Details

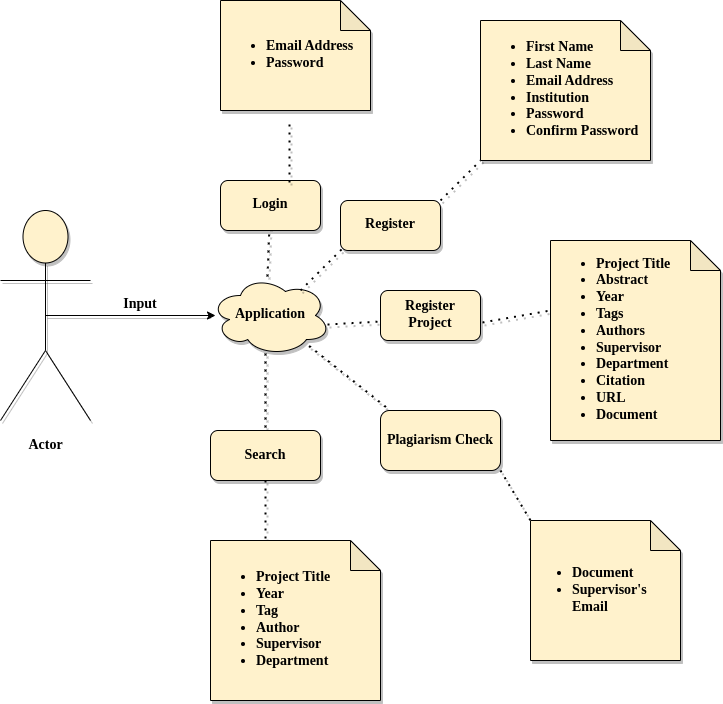
This is a web page that renders detailed information about a specific project.

#### 3.2.2.3 Search Result

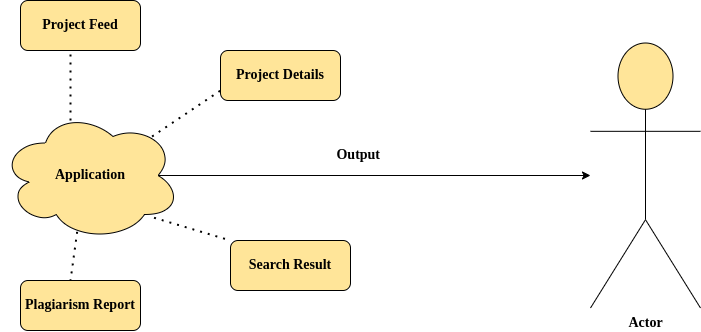
This renders a feed displaying all results gotten from the execution of a search query executed by the search field input.

#### 3.2.2.4 Plagiarism Report

This is a description of the level of plagiarism on the document, once a plagiarism check has been carried out.



**Fig 3.2.1 Input Design**



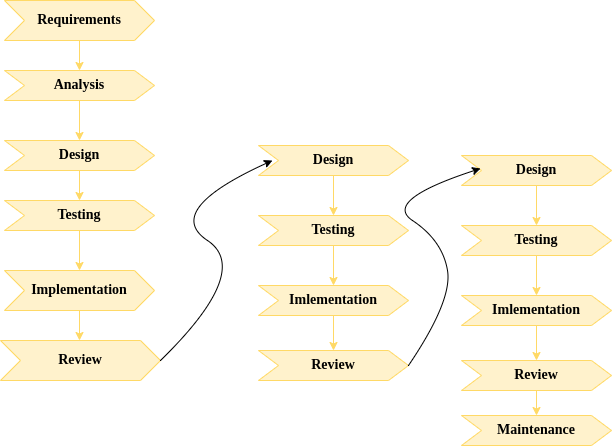
**Fig. 3.2.2 Output Design**

## 

## **3.3 SOFTWARE DEVELOPMENT PROCESS MODEL**

The model describes the steps and chronological order of a process. It is an outline of the sequence in which the process's activities are performed. The adopted process model is the iterative development model. The advantage of this model is that it provides a working model at an early stage and at every stage of development, making it easier to find design flaws. Another advantage is that it costs less to change the system requirements. There’s a better understanding of the system after every iteration. In essence, the iterative model breaks down the development process of an application into smaller parts. Once the initial planning process is completed, some other stages are repeated.

### 3.3.1 Phases of Iterative Development Process Model



#### Figure 3.3.1 Phases of Iterative Development Process Model

1. **Requirements Gathering and Analysis:** This phase of the process model involves requirements analysis whereby the requirements are gathered in detail by observation and interview to understand the user’s need for the system. The requirements are gathered at an early stage to aid the process of subsequent phases.
2. **Design:** During this phase, the team makes a preliminary design of the system. The Design stage is where many potential solutions for the project are analyzed. The solutions are narrowed down. After that, the most effective and efficient way to construct the system is decided. The team designs the software using different diagrams like UML designs.
3. **Implementation:** All the requirements and design plans are implemented and coded in this stage. This is the point in the project when the actual construction of the system starts using technological tools.
4. **Testing:** Here, several standards and norms are tested against the current build iteration to see if they are met. The purpose of these testing techniques is to identify any faults or errors in our system. The project's solution is revalidated for stability. In other words, it is made sure that fixing one bug does not introduce any more ones into our system.
5. **Review:** when all previous stages are completed. A meticulous evaluation is done on the system developed up until this stage. The development team and the stakeholders are able to examine the system and give their feedback regarding various aspects of the system.
6. **Maintenance:** Once the final system has been developed and properly tested, it is thoroughly tested and deployed. The system then undergoes routine checks and maintenance activities for reducing downtime and failure while increasing the life-span.

## **3.4 REQUIREMENTS ANALYSIS**

The requirements describing the functionality of the system are discussed below and are classified into functional and non-functional requirements.

### 3.4.1 Functional Requirements

#### 3.4.1.1 User Requirements

1. The user shall be able to register to create an account
2. The user shall be able to login to access the application
3. The user shall be able to modify their profile details
4. The user shall be able to register a project
5. The user shall be able to upload documents
6. The user shall be able to view projects
7. The user shall be able to search for projects
8. The user shall be able to download project documents
9. The user shall be able to carry out a plagiarism check
10. The user shall be able to view the report of the plagiarism check

**3.4.1.2 System Requirements**

1. The system shall allow the user create an account
2. The system shall validate the user upon login request
3. The system shall provide error messages for invalid input
4. The system shall allow the user register a project
5. The system shall allow a user view available projects
6. The system shall allow the user search for projects by title
7. The system shall allow the user search for projects by year
8. The system shall allow the user search for projects by tag
9. The system shall allow the user search for projects by author
10. The system shall allow the user search for projects by supervisor
11. The system shall carry out a plagiarism check on documents
12. The system shall generate a report after plagiarism check has been done.
13. The system shall ensure all documents are checked for plagiarism before allowing upload on the repository.

### 3.4.2 NON-FUNCTIONAL REQUIREMENTS

1. **Portability:** The system shall be be available on any device that has access to the internet
2. **Platform Independence:** The system shall run as a web application on a web browser
3. **Availability:** The system shall be a available seven days in a week, every day of the year
4. **Security:** The system shall ensure only valid passwords with minimum length of 8, alphanumeric and with symbols.
5. **Scalability:** The system shall be scalable enough to accommodate 5,000 visits at a time while ensuring optimum performance.
6. **Performance:** The system shall ensure page load is 6 seconds or less in a chrome or mozilla browser.
7. **Maintainability:** The Mean-Time-To-Restore (MTTRS) system following a system failure must not be greater than 10 minutes.
8. **Reliability:** The system shall operate without failure in 95 percent of use cases.
9. **Localization:** The system’s date format shall be as follows: month.date.year.
10. The system shall ensure documents with a plagiarized score of above 25 percent are not uploaded on the repository.

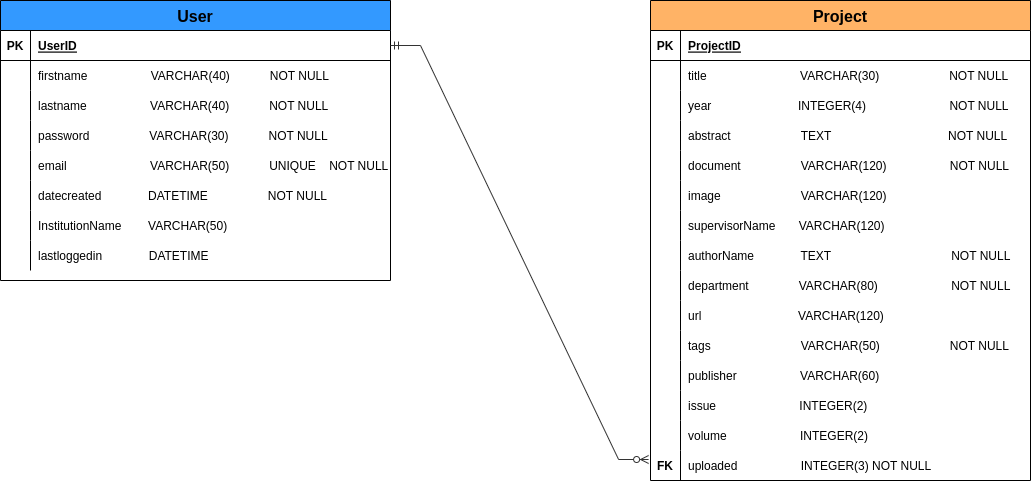
## **3.5 SYSTEM DESIGN**

### 3.5.1 USE CASE DESIGN



**Figure 3. UseCase Design**

### 3.5.2 DATABASE DESIGN

**Figure 3. Database Design**

### 

## **3.6 DESIGN AND DEVELOPMENT TOOLS**

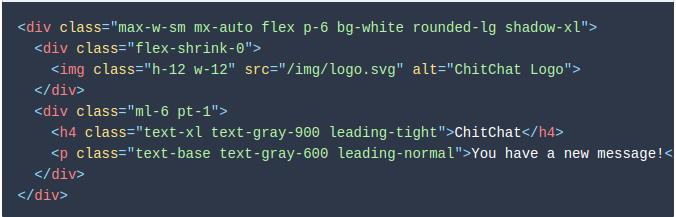
**HTML:**

The HyperText Markup Language or HTML is the standard markup language for documents designed to be displayed in a web browser. It is the code that is used to structure a web page and its content. Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML originally featured cues for the document's design and semantically explains the structure of a web page. Images and other objects, like interactive forms, may be embedded within the produced page using HTML techniques.

**Tailwind CSS:**

A utility-first CSS framework called Tailwind CSS offers a number of these opinionated, one-purpose utility classes that may be used to design elements directly inside of markup languages like html. The old method, where custom designs necessitate custom CSS, is to create CSS every time you need to style something on the web. When using Tailwind, you may create bespoke designs without creating CSS by adding pre-existing classes right to your HTML.

**Fig 3.3 Writing Custom CSS**

****

**Fig 3.3 Using Utility classes in Tailwind Css**

**Django:**

Django is a free and open-source, Python-based web framework that follows the model–template–views architectural pattern, much similar to the model-view-architectural pattern. It enables rapid development of secure and maintainable websites. It helps focus on writing an application without needing to reinvent the wheel. The tiniest distinction in all of this that might be the most perplexing is how Django recommends that the View contain the business logic in addition to the presentation logic, as opposed to just the presentation logic as in the standard MVC pattern, and that the Template should handle all of the presentation logic by itself, whereas the MVC pattern contains no Template component at all. Django's design is therefore sometimes known as the Model-Template-View + Controller when contrasted to the traditional MVC paradigm, with the Controller frequently omitted because it is already a feature of the framework.

**PostgreSql:**

PostgreSQL, is a free and open-source relational database management system emphasizing extensibility and SQL compliance. It's available on major operating systems such as [macOS](https://www.google.com/search?sxsrf=ALiCzsb6fBrKqcXpz6p5zTXsLVWfeZjwJg:1668905090314&q=macOS&stick=H4sIAAAAAAAAAONgVuLQz9U3MDWtLFrEypqbmOwfDADb5b_yFAAAAA&sa=X&ved=2ahUKEwiXrsn4w7v7AhVPPuwKHaIwBcQQmxMoAXoECG8QAw), [Windows](https://www.google.com/search?sxsrf=ALiCzsb6fBrKqcXpz6p5zTXsLVWfeZjwJg:1668905090314&q=Windows&stick=H4sIAAAAAAAAAONgVuLQz9U3MCmKt1jEyh6emZeSX14MAFTjqsQWAAAA&sa=X&ved=2ahUKEwiXrsn4w7v7AhVPPuwKHaIwBcQQmxMoAnoECG8QBA), [Linux](https://www.google.com/search?sxsrf=ALiCzsb6fBrKqcXpz6p5zTXsLVWfeZjwJg:1668905090314&q=Linux&stick=H4sIAAAAAAAAAONgVuLUz9U3SCuoqipYxMrqk5lXWgEATgerNhUAAAA&sa=X&ved=2ahUKEwiXrsn4w7v7AhVPPuwKHaIwBcQQmxMoA3oECG8QBQ). It supports text, images, sounds, and video, and is very much compatible with the python programming language, under which django is a framework.



**Visual Studio Code:**

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft with the Electron Framework, for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git for version control.

**Git**

Git is free and open source software for distributed version control: tracking changes in any set of files, usually used for coordinating work among programmers collaboratively developing source code during software development. A web application- github, based off git provides a cloud-based git repository providing easy accessibility to the codebase for collaboration between contributors.

**3.7 SYATEM EVALUATION**

**Google Pagespeed Insights**

A free performance tool called PageSpeed Insights evaluates a webpage's content and delivers page speed ratings for both the mobile and desktop versions of the website. With the help of PageSpeed Insights, you can take action to improve the functionality of your website. The Google page speed test's suggestions are based on the most recent best practices for mobile and desktop web performance. The Google PageSpeed Insights' advanced data visualization, filtering, tagging, and snapshot technologies help to provide a complete solution to improve a website's search engine ranking, amp up its performance, and improve the user experience.

**FIgure 3. Google Pagespeed Insights Logo**

# REFERENCES

[1] Iris Xie, Krystyna K. Matusiak. (2016). Chapter 9 - Digital preservation, Discover Digital Libraries, Elsevier. (Pages 255-279), ISBN 9780124171121.

[2] Robert Allan. (2009). Managing and using digital information (Pages 55-77), In Chandos Information Professional Series, Virtual Research Environments, Chandos Publishing. ISBN 9781843345626.

[3] Oladiran, M.T., Bentley, G., & Jain, P. (2013). The role of institutional repository in digital scholarly.

[4] Mishra, R. K. (2016). Definitions, Issues And Challenges Of Digital Libraries. *Innovare Journal of Education*, *4*(3), 1–3.

[5] Okon, R. , Eleberi, E. and Uka, K. (2020) A Web Based Digital Repository for Scholarly Publication. *Journal of Software Engineering and Applications*, **13**, 67-75. doi: [10.4236/jsea.2020.134005](https://doi.org/10.4236/jsea.2020.134005).

[6] Mgonzo, Wasiwasi & Yonah, Zaipuna. (2014). Design and Development of a Web Based Digital Repository for Scholarly Communication: A Case of NM-AIST Tanzania. International Journal of Knowledge Content Development & Technology. 4. 97-108. 10.5865/IJKCT.2014.4.2.097.

[7] Smith et al. (2003). DSpace: An Open Source Dynamic Digital Repository. D‐Lib Magazine, 9(1).

[8] Vangie Beal. (2022. Repository. https://www.webopedia.com/definitions/repository/

[9] Candela, Leonardo & Castelli, Donatella & Pagano, Pasquale. (2011). History, Evolution and Impact of Digital Libraries. 10.4018/978-1-60960-031-0.ch001.

[10] Eric B. (2019). A Brief History of Digital Preservation. https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1033&context=ml\_pubs

[11] Okon, R. , Eleberi, E. and Uka, K. (2020) A Web Based Digital Repository for Scholarly Publication. *Journal of Software Engineering and Applications*, **13**, 67-75. doi: [10.4236/jsea.2020.134005](https://doi.org/10.4236/jsea.2020.134005).

[12] Budapest Open Access Initiative. (2002). https://www.budapestopenaccessinitiative.org/

[13] Chan, L. (2004) Supporting and Enhancing Scholarship in the Digital Age: The Role of Open Access Institutional Repositories. Canadian Journal of Communications, 29, 277-300.[https://doi.org/10.22230/cjc.2004v29n3a1455](https://doi.org/10.22230/cjc.2004v29n3a14557-300)

[14] Lynch, C. and Lippincott, J. (2005) Institutional Repository Deployment in the United States as of Early 2005. D-Lib Magazine, 11, No. 9. https://doi.org/10.1045/september2005-lynch

[15] Crow, Raym. "The Case for Institutional Repositories: A SPARC Position Paper." ARL, no. 223 (August 2002): 1-4. .

[16] Lyle, J. (2017). Making the Case for Disciplinary Data Repositories. In Lisa Johnston (editor), Curating Research Data, Volume Two: A Handbook of Current Practice (162-164). Chicago: Association of College and Research Libraries.

[17] Patil, Sudhir S. (2019). Overview of Plagiarism Checkers and Plagiarism Detection Tools: A Study.

[18] F.E.J. Valpy (1999). Etymological dictionary of the latin language: Adegi graphics LLC

[19] Rabab A.A. Mohammed, Omar M. Shaaban, Dalia G. Mahran, Hamdi N. Attellawy, Ahmed Makhlof, Abdulkader Albasri. (2015). Plagiarism in medical scientific research, Journal of Taibah University Medical Sciences. Volume 10, Issue 1, Pages 6-11. ISSN 1658-3612.

[20] Jyotindu Debnath. (2016) Plagiarism: A silent epidemic in scientific writing – Reasons, recognition and remedies, Medical Journal Armed Forces India, Volume 72, Issue 2, Pages 164-167, ISSN 0377-1237.

[21] Adegbile, A.A., Ayobiolaja, S.P., Olanrewaju, O.T., Togun, O.A., Nwufoh, C.V. (2018), Design and implementation of a virtual project repository system. International Journal of Computer Science and Information Security (IJCSIS), Vol. 16, No. 7, July.

[22] O’Brien, K. (2019). ResearchGate. *Journal of the Medical Library Association: JMLA*, *107*(2), 284.

[23] Lex, E. (2015). A critical look at the ResearchGate score as a measure of scientific reputation.

[24] DEL ROSARIO, Marco Jr. SARENO, Julius Walailak Journal of Science & Technology. (2020) Vol. 17 Issue 7, p726-744. 19p.

[25] Shankar, M.P. & Ramasesh, C.P. Anti-plagiarism Software – A tool to ensure quality research output. In National Conference on Librarians and Librarianship in Transition: Challenges and opportunities LIBTRANS, 22-23 August, 2014, Mangalore.2014.

[26] Chinedu, Ugo & Ikerionwu, Charles & Obi, Nwokonkwo. (2020). Plagiarism Detection Systems. International Journal of Scientific and Research Publications (IJSRP). 10. p9969. 10.29322/IJSRP.10.03.2020.p9969.

[27] Shahabi, M. (2012). Comparing three plagiarism tools (Ferret, Sherlock, and Turnitin). *International Journal of Computational Linguistics*, *3*(1), 53-66.

[28] Stefania Manca (2018) ResearchGate and Academia.edu as networked socio-technical systems for scholarly communication: a literature review. Institute of Educational Technology, National Research Council of Italy, Genoa, Italy. Research in Learning Technology Vol. 26, 2018

[29] <https://medium.com/shecodeafrica/understanding-the-mvc-pattern-in-django-edda05b9f43f>

[30] https://v1.tailwindcss.com/docs/utility-first