

**JRMSU RESEARCH DEVELOPMENT AND EXTENSION PORTAL WITH
PLAGIARISM DETECTOR**

A Thesis

Presented to the Faculty of the College of Computing Studies
Jose Rizal Memorial State University
Main Campus, Dapitan City

In Partial Fulfillment
Of the Requirement for the Degree
Bachelor of Science in Computer Science

Ricalyn G. Magallon
Shelo M. Paglinawan
Ronel A. Sta Ana

December 2022



Republic of the Philippines
**JOSE RIZAL MEMORIAL STATE
UNIVERSITY**

The Premier University in Zamboanga Del Norte
Main Campus, Dapitan City



CERTIFICATION

This is to certify that (**Ricalyn G. Magallon, Shelo M. Paglinawan, Ronel A. Sta Ana**)
are taking Res 1 – Research Methodology in Computing Leading to the degree of
Bachelor of Science in Computer Science during the second semester, School Year
2022-2023.

ARMANDO T. SAGUIN JR., MSIT

Associate Dean, College of Computer Studies

This Research entitled “**JRMSU RESEARCH DEVELOPMENT AND
EXTENSION PORTAL WITH PLAGIARISM DETECTOR**” prepared and
submitted by **Ricalyn G. Magallon, Shelo Paglinawan and Ronel A. Sta Ana** has
been examined and is recommended for oral examination on (December 19, 2022).

Research Coordinator



Republic of the Philippines
JOSE RIZAL MEMORIAL STATE UNIVERSITY
 The Premier University in Zamboanga Del Norte
 Main Campus, Dapitan City



APPROVAL SHEET

This Thesis entitled “**JRMSU RESEARCH DEVELOPMENT AND EXTENSION PORTAL WITH PLAGIARISM DETECTOR**” prepared and submitted by **Ricalyn G. Magallon, Shelo M. Paglinawan, Ronel A. Sta Ana** has been reviewed and is recommended for acceptance and approval.

English Critic

Adviser

PANEL OF EXAMINERS

APPROVED by the PANEL OF EXAMINERS ON ORAL DEFENSE on with the rating
 of

Member

Member

Chairman

Vice Chairman

Accepted and approved in partial fulfillment for the subject Project Thesis leading to the degree of Bachelor of Science in Computer Science during the second semester, SY 2021-2022

ARMANDO T. SAGUIN JR., MSIT
 Dean, College of Computing Studies



Republic of the Philippines
JOSE RIZAL MEMORIAL STATE UNIVERSITY
The Premier University in Zamboanga Del Norte
Main Campus, Dapitan City



CERTIFICATION

This Thesis entitled “**JRMSU RESEARCH DEVELOPMENT AND EXTENSION PORTAL WITH PLAGIARISM DETECTOR**” has been prepared and submitted by **Ricalyn G. Magallon, Shelo M. Paglinawan** and **Ronel A. Sta Ana** in partial fulfillment of the requirements for the course: THS 101 – THESIS 1 in Computing leading to the degree of Bachelor of Science in Computer Science has been reviewed, checked and verified by the undersigned as to grammar, coherence and organization.

English Critic

DEDICATION

This study is dedicated to the family of the researchers who have expressed their unending support during the making of this research. The researchers are very grateful to the Magallon Family, Paglinawan Family, and Sta Ana Family for being present and fulfilling the needs of their children financially and emotionally.

This study is greatly dedicated to the highly respected Alma Mater of the researchers, Jose Rizal Memorial State University, which have presented them with quality education and knowledge. With this, the Academe have molded the researchers to become holistically developed individuals.

The researchers are filled with gratitude for the aid that they have received from different individuals. Their guidance, understanding, and assistance has paved the way for the researchers into this study.

Ricalyn G. Magallon
Shelo M. Paglinawan
Ronel A. Sta Ana

ACKNOWLEDGEMENT

The whole experience would not have been possible without the assistance and cooperation of the few people who have fulfilled their duties. The researchers would like recognize and convey their heartfelt gratitude to the following persons who have contributed to the success of the study.

To **Prof. Joseph Aurelius Jacinto MSIT** the adviser, for dedicating his time and knowledge for the fulfillment of the study;

To the Faculty and staff of the College of Computing Studies, for their encouraging comments and support;

To the parents of the researchers, for showing their love and support to this whole journey financially and morally;

To their friends and classmates, for being great helping hands during the study.

Lastly, to the Almighty Father, who have bestowed the researchers His eternal protection, guidance, and blessings.

*Ricalyn G. Magallon
Shelo M. Paglinawan
Ronel A. Sta Ana*

TABLE OF CONTENTS

	Page
Title Page	i
Approval Sheet	ii
Dedication	iii
Acknowledgement	iv
List of Tables	viii
List of Figures	ix
Chapter	
1 THE PROBLEM AND ITS SCOPE	
Introduction	1
Statement of the Problem	3
Significance of the Study	4
Scope and Delimitation of the Study	5
Definition of Terms	6
Theoretical and Conceptual Framework	8
2 REVIEW OF LITERATURE AND STUDIES	
Literature	13
Foreign	19
Local Studies	22
3 RESEARCH METHODOLOGY	
The Design	27

Research Setting	27
Respondents of the Study	29
Research Instrument	30
Validation of Instrument	30
Scoring Procedure	31
Data Gathering Procedure	32
Statistical Treatment of Data	33
Project Development Process	33
Requirement Analysis	33
System Designing	34
System Architectural Design	35
Hierarchical Input Output Process	36
Use Case Diagram	37
Class Diagram	38
Sequence Diagram	40
Swimlane Activity Diagram	42
Entity Relationship Diagram	43
Data Flow Diagram	44
Coding	53
Testing and Evaluation	54
System Testing Procedure	55
System Evaluation	55
System Implementation	56

REFERENCES**APPENDICES**

LIST OF TABLES

Table		Page
1	Respondents Distribution	33
2	Cronbach's Alpha	36
3	Scoring Procedure	36
4	Minimum Software & Hardware Requirements	39
5	The Profile of Respondent in terms of Profession, Gender and Age Survey	64
6	The Extension Portal user Evaluation for Functionality Contains a Plagiarism detection.	65
7	The Extension Portal user Evaluation for Reliability Contains a Plagiarism detection	66
8	The Extension Portal user Evaluation for Usability Contains a Plagiarism detection.	67
9	The Extension Portal user Evaluation for Efficiency Contains a Plagiarism detection.	68
10	The Extension Portal user Evaluation for Maintainability and Probability Contains a Plagiarism detection	69
11	Average Weighted Mean	70

LIST OF FIGURES

Figure		Page
1	Conceptual Framework	12
2	Schema of the Study	14
3	Map Location of six (6) JRMSU Campuses	32
4	System Architectural Design	41
5	Hierarchical Input Process Output (HIPO)	42
6	Use Case Diagram	43
7	Class Diagram	45
8	Sequence Diagram	47
9	Swimlane Activity Diagram	49
10	Entity Relationship Diagram	51
11	Data Diagram	53

Chapter 1

The Problem and Its Scope

Abstract

In this century, technology has become a part of daily lifestyle. Technology is the application of scientific knowledge to the practical goals of human life, or, as it is sometimes referred to, to the modification and manipulation of the human environment. The connection between humans and technology is the one reason for economic development. Technology makes individuals' lives better. It has evolved into a primary driving force in the advancement of society's needs. Humans use technology for various purposes. Technology is human knowledge; it includes tools, materials, and systems. Technology has become a hobby to somebody in the way of communicating, buying things, learning, training programs, data security, grading systems, and so much more. Humans benefit from technology when it is used correctly, but the opposite is true when it is used maliciously. Technology will always evolve with the advancement of the human race, which influences many factors in the future.

This study aims to address the problem of improving the effectiveness of plagiarism detection methods by integrating various analysis methods using machine learning. Despite significant advances in the field of plagiarism detection, including better semantic text analysis and investigation of non-textual content features, there is still a need for more thorough performance evaluations of plagiarism detection systems. Thus, the study aims to identify the research gap and propose new directions for future research to enhance the accuracy and efficiency of plagiarism detection methods. (Foltýnek 2019).

The problem identified in this study is the need for more effective methods for preventing plagiarism and evaluating the effectiveness of current strategies and tools. The study highlights that while there is an emphasis on improving researcher awareness about the negative impact of plagiarism on scholarly publications, there is a lack of specific guidance on how to address the root causes of plagiarism such as poor academic English skills and a lack of creative thinking. Moreover, the study identifies a research gap in evaluating the effectiveness of current strategies and tools for plagiarism prevention and detection and suggests the need for new approaches or solutions to these issues. Therefore, the study aims to fill this research gap by proposing new ideas and methods for preventing plagiarism and evaluating the effectiveness of existing strategies and tools. (O Zimba 2021)

Plagiarism detectors are essential in maintaining the integrity of research papers. It is highly recommended to incorporate them in schools and offices to prevent any forms of plagiarism. The JRMSU system's plagiarism detector is an advanced portal that can efficiently scan and analyze the entirety of a research paper's chapters. It also provides relevant details about the paper's status, including the proposed date, start date, and completion date. The portal equipped with this tool is highly valuable for employees in the research building, as it enables them to monitor the research's progress and promptly identify any potential cases of plagiarism.

To address the research gap in the previous statement, future studies should focus on exploring new approaches to improving researcher awareness of plagiarism, including interventions aimed at developing creative thinking skills and enhancing academic English proficiency. In addition, researchers can investigate the limitations of existing plagiarism

detection software and explore new methods for detecting plagiarism, such as incorporating machine learning and natural language processing techniques. Furthermore, a rigorous evaluation of the effectiveness of current plagiarism prevention and detection strategies is necessary. This can involve large-scale testing and comparison of various tools and techniques. Conducting such research can result in the development of more effective and efficient methods for preventing and detecting plagiarism in scholarly publications.

This paper proposes a solution to overcome the limitations of portals with plagiarism detectors. Many educational institutions use streamlined systems, and the plagiarism detector in such systems offers several advantages, including privacy, integration with Microsoft products, compatibility with multiple document formats, efficient performance, advanced reporting capabilities, and useful functionalities for detecting plagiarism. The plagiarism detector is crucial for identifying plagiarism in research papers and monitoring their status. The system is unique and serves as a valuable resource for detecting instances of plagiarism in academic writing.

This study aims to provide a comprehensive understanding of the importance of plagiarism detection in the research process and to equip researchers with the necessary skills and tools to avoid ethical misconduct in their scholarly publications. By utilizing an advanced and modern portal with a plagiarism detector, researchers can ensure that their work meets the highest standards of academic integrity and credibility. Additionally, the results of this study can be used to inform policy decisions and institutional practices related to research and publication ethics at the JRMSU System. Ultimately, this study contributes to the advancement of knowledge and the promotion of ethical research practices within the academic community.

Statement of the Problem

This study aims to implement a Portal with Plagiarism Detector in University Research Development and Extension for Jose Rizal Memorial State University – Dapitan Campus during school year 2022-2023.

Specifically, this study sought to answer the following question:

1. What is the current portal of JRMSU for Research Development and Extension as well as Plagiarism system?
2. How does the plagiarism checker work in the JRMSU Research Development and Extension Portal?
3. What other similar system features can be fused to come up with a hybrid or the proposed system?
4. How can the system be assessed in terms of:
 - 4.1 Functionality;
 - 4.2 Reliability;
 - 4.3 Usability;
 - 4.4 Efficiency;
 - 4.5 Maintainability; and
 - 4.6 Portability?

Significance of the Study

This newly developed JRMSU Research Development and Extension Portal with Plagiarism Detector is available exclusively on all campuses of the JRMSU system. It is important for the university to have its own plagiarism detection system portal, since students are required to submit researches or thesis as a course requirement. Also, faculties are required to submit researches every semester. Because of these, using an online plagiarism detection have been a burden due to its subscription fees and besides, although there are free online plagiarism system, the number of words and paragraphs which can be scanned are limited on free subscription or account.

The purpose of this study is to create a system through the combination of a portal and a plagiarism detector. The specific goal of this system is to: (1) determine the status of submitted research paper in JRMSU (2) Detect plagiarized documents. (3) Prevent plagiarism cases in the JRMSU system. (4) Provide a free and better plagiarism detection system for JRMSU.

Researchers in JRMSU- Since this is online system software, researchers can submit the partial research paper at any time and from any location by using an internet connection.

Employee – Ease to manage research document access on the employee side.

Administrator – Since this is online system software, the administrator can also manage it with ease.

Future Researcher's – The result of this study will serve as a guide and resource for aspiring researchers.

Scope and Delimitation of the Study

This research focused on the portal with the plagiarism detector. This system is specifically designed to help Research Development and Extension building employees and JRMSU system researchers track the status of research papers and detect plagiarism. This system is limited to the researchers in the JRMSU system and the VPRDE employees. The user accesses the system using the default username and password provided by the administrator which the administrator can manage anytime. This system is accessible from a PC, Mac, or mobile phone.

This system was developed especially for the Research Development and Extension building to have an advanced and high-tech portal with a plagiarism detector. The system is a combination of two interfaces: (1) Portal, which represents the body of the system and can perform tasks including user log-in, placement of data, personal information management, data security, and managing research status. (2) Plagiarism Detector: Based on the number of words per minute, it can detect plagiarized documents in research papers uploaded to the portal. This study will be conducted on all branches of the JRMSU system. The respondents were active JRMSU system researchers, VPRDE employees, and instructors who covered research and extension development.

In addition, the system designed to have several capabilities related to research and presentation management. Firstly, it should be able to generate a list of available research or extension documents that are relevant to the author. This list should be filterable by different categories such as research, extension, publication, or conference paper. Additionally, the system should provide search functionality to help the author find specific documents based on title, author, date, or keyword. The system should display information about each document, such as the document type, publication date, author name(s), and a

brief summary of the content. Furthermore, the author should be able to download or email a selected document to themselves or others. The system should also provide a feature for the author to rate or review the document for other users to see.

Regarding presentation management, the system provide a template for the presentation schedule form that includes fields such as the presentation title, date, time, location, and duration. It should allow the author to select a presentation category, such as research, extension, or poster session. The author should also be able to invite other presenters or co-authors to join the presentation. The system should provide a feature for the author to submit the presentation schedule form electronically or print it out for manual submission. Finally, the system should send a confirmation email to the author once the presentation schedule form is received and approved.

After the presentation, the system provide a checklist of compliance requirements that the author needs to meet, such as submitting a copy of the presentation slides, filling out a survey or evaluation form, or submitting a final report. The system should provide a deadline for when the author needs to complete each requirement and send reminder emails as the deadline approaches. The author should be able to submit the compliance requirements electronically or manually, and the system should provide a confirmation email once the compliance requirements have been met.

Access to the internet is crucial for the applicability of this study as it involves the use of an online software system to execute commands and perform the interfaces of the Portal with Plagiarism Detector. Without internet access, it would not be possible to utilize this system effectively, rendering the study not applicable.

Definition of Terms

To facilitate the understanding of this study, different terms are defined herein.

API Key – A set of functions and procedures allowing the creation of applications that access the features or data of an operating system, application, or other service.

Data Base - A structured set of data held in a computer, especially one that is accessible in various ways.

Detector - A device or instrument designed to detect the presence of a particular object or substance and to emit a signal in response.

Efficiency – A set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions.

Local Host – This refers to standards host name given to address of the loopback network interface.

Maintainability - A set of attributes that bear on the effort needed to make specified modifications.

MySQL – This refers to a relational database management system (RDBMS) that runs as a server providing multi-user access to a number of databases.

Plagiarism – The practice of taking someone else's work or ideas and passing them off as one's own.

Portability – A set of attributes that bear on the ability of software to be transferred from one environment to another.

Portal - A website or web page providing access or links to other sites.

System – The used of this study is to give the complete methods in order to operate the same procedure to the given data..

VPRDE – The office of the Vice President for Research, Development, and Extension is on the JRMSU main campus in Dapitan City.

Website – Is a set of related web pages containing content such a text, images, video, audio, etc., A website is hosted on at least one web server.

XAMMP – This refers to a free and open sources cross-platforms web server solution stack package, consisting mainly of the Apache HTTP Server.

Operational and Definition of Terms

Portal with Plagiarism Detector can prevent plagiarism activities in JRMSU system.

Reliability. The reliability is important because it measures the quality of the research and findings that are true or accurate from a research study are often reliable. As well as we will be ensured in every component of the system is reliable and works in optimal portal with plagiarism detector.

Functionality. In information technology, it is the authentication and identification can do for user.

Usability. It is sometimes seen as purely and end user issue or something attribute that makes a product easy to understand, easy to learn, easy to use, and attractive to users. Software provider can enhance the usability of security features by integrating how user

intuitively visualize implementing the security of the system. Usable security also covers the design, development, configuration and maintenance of the tools and systems the business runs on.

Theoretical and Conceptual Framework

Plagiarism detection is a critical element of academic writing and research that aims to identify instances where a writer has used someone else's work without proper attribution.

This plagiarism detector is based on the concept of similarity in text. This theory assumes that a plagiarized document will have significant similarities with other existing documents, especially in terms of language and structure. The plagiarism detector uses various algorithms to analyze the input text and compare it with a large database of pre-existing documents. The algorithms consider factors such as word choice, sentence structure, and even formatting to determine the similarity score between the input text and the database documents.

In recent years, advancements in machine learning and natural language processing have led to the development of more sophisticated plagiarism detection theories that can account for nuances in language and detect more subtle forms of plagiarism. These theories typically use a combination of algorithms and machine learning models to analyze text and detect patterns of similarity.

This study revolves around two interfaces: (1) Portal and (2) Plagiarism Detector together with an aspect such as functionality, reliability, usability, efficiency, maintainability, and portability.

The portal represent as the body which represents the body of the system and can perform tasks including user log-in, placement of data, personal information management, data security, and managing research status.

The plagiarism detector is capable of detecting plagiarized documents in research papers uploaded to the portal.

This system's portability allows a system to move across environments rather than just platforms. To clarify, a computer platform generally refers to the operating system and computer hardware only. Since this system executes using an internet connection, users can access the system anytime and anywhere.

Figure 1 is a conceptual framework, which is a structure that outlines the key concepts, variables, and relationships involved in a research study. It provides a broad overview of the research problem and serves as a guide for developing research questions, hypotheses, and data collection methods. The statement outlines the process flow of a system that involves a portal, a database, and a plagiarism detector. Authors upload documents to the portal, which stores them in a database for further processing. The plagiarism detector then analyzes the documents by comparing them to a database of other texts to detect any instances of plagiarism. Once the analysis is complete, the system generates a report, which is then sent to the user through the portal interface. The report provides feedback on the degree of similarity between the uploaded document and other texts and identifies any matching sources. The goal of the system is to promote originality and ensure proper attribution of sources.

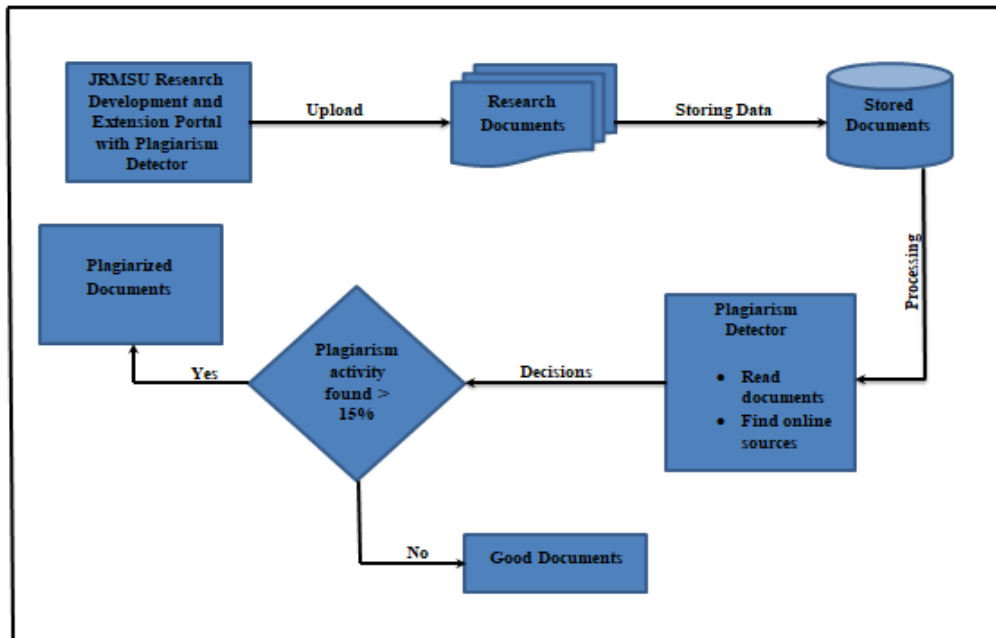


Figure 1: Conceptual Framework

As illustrated in Figure 2. The input box contains Computer, Mobile Phone, Internet Connection, Network Component, and PHP Programming Language which serve as the foundation for creating the output. Internet connection is the requirements to give functions the Network Components such as clients, servers, channels, interfaces devices, and operating system.

The process uses the Software Development Life Cycle (SDLC) as its model for development, it has seven stages, respectively; Data Gathering – to seek information of approaches in implementing the Portal with Plagiarism Detector – identifying the hardware and software minimum requirements, Designing – the steps in designing the approach for implementing the Portal with Plagiarism Detector, Coding – data converted to computer commands and instructions using PHP programming language, Testing and Evaluation – checking the accuracy and speed of the computerized procedure, Implementation – current

system is executed for a period of time as comparison for a possible new system, Documentation – software documentation for the developed software. The project method used is an Incremental Build Model where the product is designed, implemented and tested incrementally until completion; adding a little more each time. As games have many intricate parts and scale very quickly, a modular approach is heavily beneficial for one.

The feedback loop includes several aspects. Firstly, if there is a lost internet connection, it can greatly affect the functionality of the portal with the plagiarism detector. This highlights the importance of a stable internet connection to ensure that the system works effectively.

Secondly, the feedback loop may identify features that need to be modified, changed, or added to improve the functionality of the system. This could involve adding new plagiarism detection algorithms, improving the user interface, or enhancing the system's reporting capabilities.

Thirdly, the feedback loop may suggest improvements or changes to the code that underpins the portals plagiarism detector system. This could involve optimizing the algorithms used by the plagiarism detector, improving the efficiency of the system, or making it more secure.

Lastly, the feedback loop may recommend code refactors for the portals plagiarism detection system. This could involve improving the system's functionality, reliability, usability, efficiency, maintainability, and portability. These are all important aspects of a well-designed system that operates effectively over the long term.

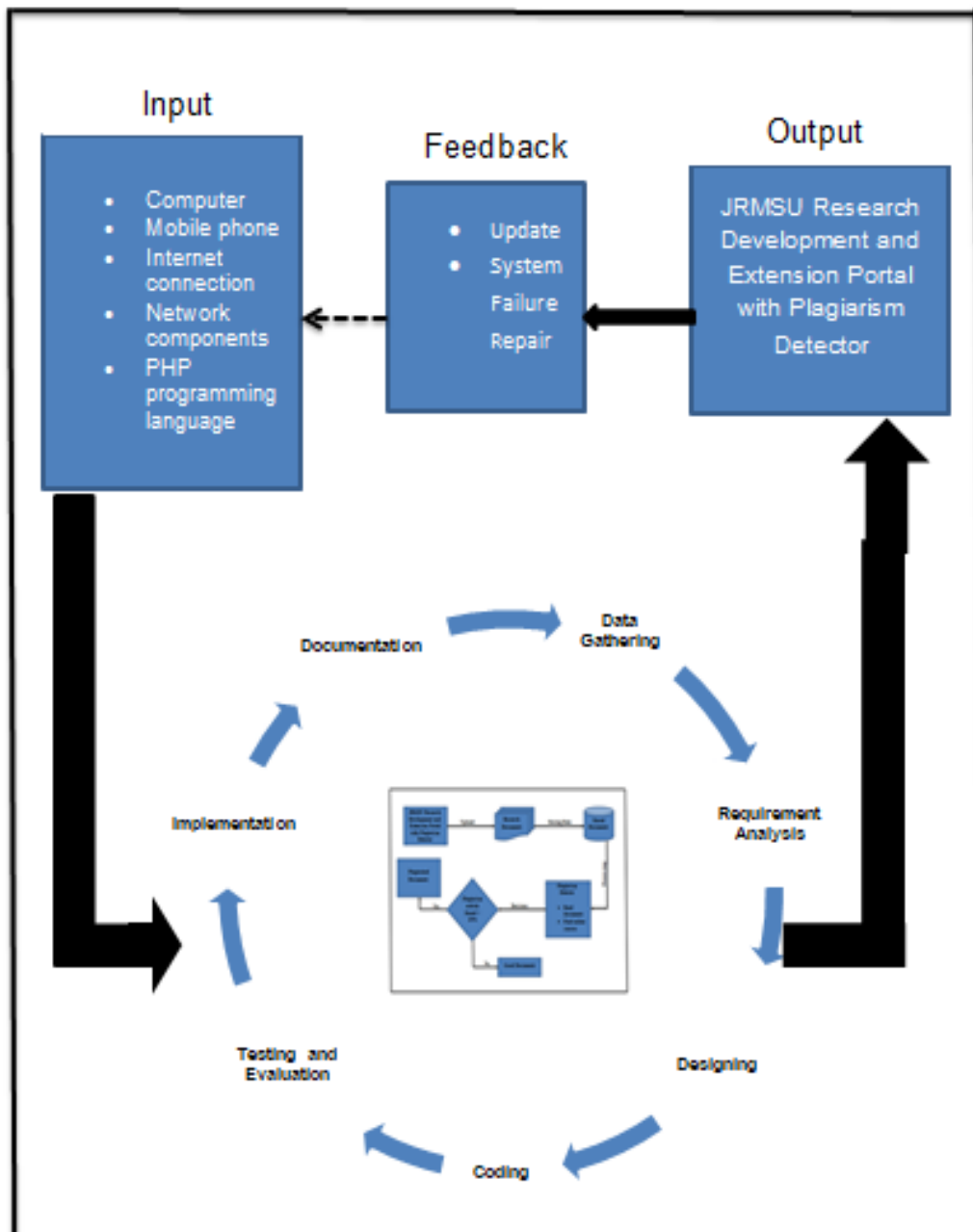


Figure 2: Schema of the Study

Chapter 2

Related Review of Literature

This chapter presents different works and studies which almost have the same concepts and theory on Portal with Plagiarism Detector. These are mostly collected from the internet, which has provided the researchers with the ideas and principles in the development of the study.

Literature

Technology advancement has paved the way for publishing internet articles across many institutions, offices, and throughout the internet in the past year, abandoning the outdated approach. Although plagiarism is not a criminal or civil offense, it is illegal if it infringes on an author's intellectual property rights, including copyright or trademark.

In 2019, T. Foltnek, N. Meuscke, and B. Gipp state the survey results. There are 239 research papers published between 2013 and 2018 that have been detected as containing academic plagiarism. To structure the presentation of the research contributions, the researcher proposes novel technically oriented typologies for plagiarism prevention and detection efforts, the forms of academic plagiarism, and computational plagiarism detection methods. The researcher shows that academic plagiarism detection is a highly active research field. Over the review period, the field has seen major advances regarding the automated detection of strongly obfuscated and thus hard-to-identify forms of academic plagiarism. "This is proof that the plagiarism detector is an effective method for quickly detecting plagiarized documents." "It is better to research methods to reduce this case and preserve dignity." T. Eisa, N. Salim, and S.

Alzahrani (2015) focused on analyzing the state-of-the-art techniques used to detect plagiarism in terms of the limitations, features, taxonomies, and processes. To prevent stealing, the document will be run through a plagiarism detector to see if it has been copied from the internet without the reference. A plagiarism detector will detect it before posting it on the internet for online documentation, either to avoid embarrassing yourself in front of others or the school, or both. Some cases of plagiarism happen inside the school. The researcher developed a visual analysis tool to support the verification, assessment, and presentation of alleged cases of plagiarism.

(P. Riehmann, M. Potthast, B. Stein, and B. Froehlich, 2015) said that. "The researcher developed a visual analysis tool to support the verification, assessment, and presentation of alleged cases of plagiarism." The analysis of a suspicious document typically results in a compilation of categorized "finding spots." The different views are tied together by versatile navigation and selection operations. "Technology is a big help in such cases because if only relying on the manual method, the process will take longer, and there is also a high possibility that everything will not be read."

"During system implementation, the researcher considered various factors and conducted analyses for system compatibility." A plagiarism detection system is a software that offers the necessary data and information to support a human reviewer in asserting whether a document, or parts of it, are an instance of plagiarism, as well as the severity of the case and the sources from which it was copied (M. Abdelhamid, S. Batata, and F. Azouaou, 2022).

"In this study, the researcher employs the stylometry method to assess the possibility of plagiarism." M. Sallal and R. Iqbal (2019) V. Palade S. Amin and V. Chang

said that the methods of tackling plagiarism were stimulated by the authorship analysis approaches, which use several text analysis techniques to infer the authorship of suspicious texts. This method targets the authorial attributes and ignores any content-related topics. The core component of this method is stylometry, which relies on deriving sets of features based on MCW frequencies. The performance of this method will be measured based on how the derived sets of features perform using MLPs and other machine learning algorithms.

"Plagiarism detection as a system provides a greater percentage of the correct result or outcome." A. S. Bin-Habtoor and M. A. Zaher (2012) said that detecting plagiarism is important to judge and mark students' work, especially for postgraduates who are strictly prohibited from cheating, rewording, rephrasing, or restating without referencing. In this regard, numerous plagiarism detection systems have been developed. The system can be classified into two main categories: web-enabled systems and stand-alone systems.

(Salim Razi 2015) said that plagiarism is defined as "the practice of claiming credit for the words, ideas, and concepts of others" (American Psychological Association [APA], 2010, p. 171). The challenges caused by plagiarism are becoming more important in parallel with developments in Internet technology. In general, plagiarism may occur in any aspect of daily life, such as academic studies, computer games, journalism, literature, music, the arts, politics, and many more. Unsurprisingly, higher-profile plagiarizers receive more attention from the public (Sousa-Silva, 2014). "Plagiarism causes widespread problems in society; there have been cases involving celebrities, politicians, vloggers, researchers, and other well-known individuals."

"By making a plagiarism detector, researchers will consider good qualities such as functionality, flexibility, reliability, and maintainability." (Steven Bradley 2020) said that when providing a more divergent assessment, automated plagiarism detectors are much more effective because the likelihood of false positives is much lower than in more constrained, convergent assessments. The researcher also shows how to combine this with partial automation of assessments.

(Atadjanov, J., & Atadjanov, B. 2019) used CLAD (Cross-Language Analog Detector) between test documents and indexed documents. The main difference between this method and previous versions is that it detects plagiarism in multiple languages rather than just two. While translating terms, it used the dictionary-based machine-translation method. The purpose of displaying synonym forms of terms above is to explain the comparison process to the reader when considering synonymity. A common mode of dealing with this is to apply plagiarism detector software to code submissions to check for suspected plagiarism based on how similar the submissions are.

(Deng, P., Linsky, C., & Wright, M. 2020) said that a software plagiarism detector (SPD) was developed using the Intel Pin instrumentation tool and that tested this approach on different small executable single-threaded and multi-threaded files. The results portray the validity of this method. The researcher also discusses the possibility of expanding this method to larger software applications.

(Chavan, H., Taufik, M., Kadave, R., & Chandra, N. 2021) pointed out that related studies Plagiarism, as a crucial offense, especially in academia, is not only a well-known problem in the text but is also becoming widespread in images. In this work, the performance of manifold ranking, known as the robust method among semi-supervised

methods, has been investigated by using twelve different features. As its high performance is attributed to the quality of the constructed graph, researchers applied a robust k-regular nearest neighbor (k-RNN) graph in the framework of manifold-ranking-based retrieval.

(Yalcin, K., Cicekli, I., & Ercan, G., 2022) propose that the aim of this paper is to present an automatic plagiarism detection system to identify plagiarized passages of documents. This plagiarism detection system uses both syntactic and semantic similarities to identify plagiarized passages. The researcher's proposed method is a novel contribution because of its usage of part-of-speech tag n-grams (POSNG), which are able to show syntactic similarities between the source and suspicious sentences. Each source document is indexed according to part-of-speech (POS) tag n-grams by a search engine in order to rapidly access sentences that are possible plagiarism candidates. Even though the plagiarism detection system obtains very good results just using POS tag n-grams, its performance is further improved with the use of semantic similarities. The proposed method is a novel contribution because of its usage of part-of-speech tag n-grams (POSNG), which are able to show syntactic similarities between the source and suspicious sentences. Each source document is indexed according to part-of-speech (POS) tag n-grams by a search engine in order to rapidly access sentences that are possible plagiarism candidates.

(Sibomana, E., Ndayambaje, I., & Uwambayinema, E. 2018) proposed that in the academic arena, plagiarism is on the rise. As a result, higher education institutions are enacting harsh punitive measures. Most of these measures are more reactive than preventive because the focus on plagiarism detection and the kinds of punishments to administer to students who plagiarize. In fact, these measures fail to address particular or

primary reasons for plagiarism in students' works. This could be one of the reasons why such measures have done little to reduce plagiarism.

(Bouarara, H. A., Hamou, R. M., & Rahmani, A. 2017) said that the number of plagiarism cases has increased and has become a topical problem in the modern scientific world, caused by the quantity of textual information available online and offline. The authors' work deals with the development of a new plagiarism detector system called BHA2, which has as input the suspicious text (to be analyzed) and the original texts (the learning basis). It can detect different forms of plagiarism based on: the Google API to detect cases of plagiarism with translation; text summarization to detect the plagiarism of ideas; conceptual transform to detect the plagiarism with synonymy; a bag of ashes to detect the paraphraser plagiarism; the social worker bees algorithm that was inspired by the lifestyle of social worker bees (forager, guardian, and cleaner) to select the documents' source of plagiarism; the output of the authors' system is the plagiarized passages (the copied parts from the original texts) and the plagiarism percentage for each suspicious text.

Foreign Studies

M. Abdelhamid, F. Azouaou, and S. Batata (2022) said that in academia, plagiarism is certainly not an emerging concern, but it became of a greater magnitude with the popularization of the Internet and the ease of access to a worldwide source of content, rendering human-only intervention insufficient. Despite that, plagiarism is far from being an unaddressed problem, as computer-assisted plagiarism detection is currently an active area of research that falls within the fields of information retrieval (IR) and natural language processing (NLP). Many software solutions emerged to help fulfill this task, and this paper presents an overview of plagiarism detection systems for use in Arabic, French,

and English academic and educational settings. The comparison was held between eight systems and performed with respect to the features, usability, and technical aspects, as well as the performance in detecting three levels of obfuscation from different sources: verbatim, paraphrase, and cross-language plagiarism. An in-depth examination of technical forms of plagiarism was also performed in the context of this study. In addition, a survey of plagiarism typologies and classifications proposed by different authors is provided.

G. Hu and X. Sun (2017) pointed out that plagiarism is a major problem for universities worldwide and has been a constant cause of concern in higher education. Previous research has focused on Chinese students' attitudes toward, knowledge of, and engagement in plagiarism in Chinese and overseas educational contexts, and there is also a growing body of research on Chinese teachers' understanding of and stance on plagiaristic practices. However, little research attention has been given to institutional policies on plagiarism in the Chinese context, though similar research has been conducted in other settings. This paper reports on a study that examined plagiarism policies made publicly available by eight major universities of foreign studies in mainland China. Both the structure and content of these universities' policy documents are analyzed to identify institutional understandings of, attitudes toward, and sanctions on plagiarism. The analysis reveals that despite inter-institutional variations, the policy documents are dominated by moralistic and regulatory discourses and are characterized by a conspicuous lack of an educational approach to plagiarism. Researchers argue that such an institutional approach to plagiarism is unlikely to be effective because it largely fails to support students' acquisition of academic literacy and legitimate intertextual practices.

O Zimba, A Gasparyan (2021) said that plagiarism is an ethical misconduct that affects the quality, readability, and trustworthiness of scholarly publications. Improving researcher awareness of plagiarism of words, ideas, and graphics is essential for avoiding unacceptable writing practices. Global editorial associations have publicized the statements on strategies to clean literature from redundant, stolen, and misleading information. When updating author instructions and warning plagiarists of academic and other consequences of the unethical behavior, it is recommended to consult related documents. A lack of creative thinking and poor academic English skills are believed to compound most instances of redundant and "copy-and-paste" writing. Plagiarism detection software largely relies on reporting textual similarities. However, manual checks are required to reveal inappropriate referencing, copyright violations, and substandard English writing.

AY Gasparyan and B. Nurmashev (2017) said that plagiarism may take place in any scientific journal despite currently employed anti-plagiarism tools. The absence of widely accepted definitions of research misconduct and reliance solely on similarity checks do not allow journal editors to prevent most complex cases of recycling scientific information and wasteful, or "predatory," publishing. This article analyzes Scopus-based publication activity and evidence on poor writing, a lack of related training, emerging anti-plagiarism strategies, and new forms of massive wasting of resources by publishing largely recycled items that evade the "red flags" of similarity checks. In some non-Anglophone countries, "copy-and-paste" writing still plagues pre- and postgraduate education. Poor research management, the absence of courses on publication ethics, and limited access to quality sources confound plagiarism as a cross-cultural and multidisciplinary phenomenon. Over the past decade, the advent of anti-plagiarism software checks has helped uncover

elementary forms of textual recycling across journals. But such a tool alone proves inefficient for preventing complex forms of plagiarism. Recent mass retractions of plagiarized articles by reputable open-access journals point to the critical deficiencies of current anti-plagiarism software, which does not recognize manipulative paraphrasing and editing. Manipulative editing also finds its way into predatory journals, ignoring publication ethics and allowing for nonsense plagiarism. The evolving preventive strategies are increasingly relying on intelligent (semantic) digital technologies, comprehensively evaluating texts, keywords, graphics, and reference lists. It is the right time to enforce adherence to global editorial guidance and implement a comprehensive anti-plagiarism strategy by helping all stakeholders in scholarly communication.

G. Hu, J. Lei, and S. Système (2016) said that research on plagiarism has largely left English as a Foreign Language (EFL) teachers out of the picture. This study set out to bridge the gap by comparing how 142 Chinese university EFL teachers and 270 undergraduate students viewed exemplars of unacknowledged copying and unattributed paraphrasing, two forms of intertextuality generally regarded as plagiarism in Anglo-American academia. More than half of the teacher participants had overseas academic experience. Quantitative and qualitative analyses found that the participants, though the understanding of plagiarism in English academic writing differently from Anglo-American academia, clearly disapproved of recognized cases of plagiarism. The analyses also revealed that greater knowledge of and harsher stances on both types of transgressive intertextuality were associated with wider exposure to and more experience in English academic writing. Furthermore, the participants had more similar understandings of unacknowledged copying than of unattributed paraphrasing and took harsher stances on

the former. These findings highlight complex and nuanced understandings of plagiarism and point to the crucial role of academic socialization in shaping knowledge of and attitudes toward plagiarism.

Local Studies

R. Ebarido (2018) said that higher education is the last bastion in the fight against cyber-plagiarism. Students are trained to be ethical professionals within the university context by molding the young minds through a combination of pedagogical strategies and academic programs to preserve the morals, values, and character. While technology is a catalyst for enhanced learning, it has also been used to commit illicit forms of academic dishonesty such as cyber-plagiarism. This paper applies the Theory of Planned Behavior and is complemented by a dimension borrowed from the Big Five Personality Theory to investigate the drivers of cyber-plagiarism in a university in the Philippines. A total of 309 responses were gathered to test the applicability of the proposed research model. A partial least squares structural equation model is applied to validate the research model in this study. Although ICT attitudes and peer pressure influence behavioral intentions to commit cyber-plagiarism at a certain level, this cannot be supported in this investigation. On the other hand, ICT literacy and neuroticism predict behavioral intentions to commit cyber-plagiarism, which subsequently predict the performance of such an act. Technological and pedagogical recommendations are discussed.

DEL Marco Jr., J. Sareno (2020) pointed out that in education, students attempt to copy previous works and are relying on prepared solutions available on the Internet in order to meet the requirements. This action leads to plagiarism, which is becoming part of educational institutions' concern to reduce growing academic dishonesty. With regards to

the aforementioned issue, this study aims to design and develop a plagiarism checker capable of registering documents, granting access to users, and calculating the similarity between documents. Thus, the software was constructed using HTML, PHP, JavaScript, CSS, and MySQL. The developed system is composed of three main modules: the document search, which enables users to browse documents; the document registration, which enables the administrator to add and manage the stored documents; and the document comparison, which serves as the system's plagiarism detection mechanism. The algorithm Normalized Compression Distance was used to measure similarity and the Boyer-Moore Algorithm to highlight the suspected plagiarized document. Moreover, tests were conducted to determine if the system is functioning as expected and to measure the accuracy of the output produced by the system. The developed system was evaluated using the ISO 25010 software quality model in terms of product quality and was rated by one hundred respondents. The system obtained a mean of 4.70, which is equivalent to "excellent" in descriptive terms. This validates that the objectives of the study were met and achieved. This further indicates that the system was developed according to its desired functions and requirements.

AG Roman (2018) said that plagiarism has been an issue in higher education institutions all across the globe. The prevalence of this issue has been reported in numerous studies and literature. However, only a few are reported to have conducted a study on the prevention of this case. Hence, this study was conducted aiming to minimize the occurrence of plagiarism among teacher education students to increase academic integrity and honesty by utilizing plagiarism detector software (Turnitin). Sixty-one (61) teacher education students were selected at an average effect size of 0.5 and subjected to two weeks of

treatment. One sample pre-test-post-test design of a quasi-experiment was utilized. Data were gathered using the plagiarism software (Turnitin) on the submitted research manuscripts of the students and compared the similarity indices of the papers before and after the utilization of Turnitin using a paired T-test statistical treatment at a 5% alpha level of significance. Results showed a significant decrease in the similarity index of the submitted manuscripts of the teacher education students before and after the utilization of Turnitin. Through triangulation, several factors were identified that influence students to commit plagiarism, such as self-resourcefulness, academic culture, and a weak plagiarism detection scheme. Finally, the paper gives a set of recommendations to further minimize the incidence of plagiarism.

According to R. Bonifacio (2020), plagiarism is an academic crime that involves stealing the ideas of others without giving proper credit. This study aims to identify the dominant type of plagiarism, determine the percentage of plagiarism committed by CMU undergraduate students using Grammarly, find the reasons why students plagiarize, and examine the effects of the Plagiarism Intervention Program. The study employed content analysis, interviews, and focus group discussion to substantiate the results. The analysis of the 328 samples shows that copying without citation is the dominant type of plagiarism, followed by unintentional plagiarism, patchwork plagiarism, and copying with citation. Furthermore, the overall mean percentage of plagiarism is 8.28%. Moreover, 15 out of 328 samples do not show any plagiarism. The students plagiarize because of time constraints, a lack of knowledge, and confidence in not getting caught. The incidence of plagiarism was eliminated in the second phase of the plagiarism intervention program. The

implications of the study led the researcher to propose the application of the plagiarism intervention program in the writing courses to eliminate the incidence of plagiarism.

CC Orlanda-Ventayen (2019) said that the English language has been the medium of instruction in most universities and colleges in the world and is common and mostly used to threaten the integrity of the scientific literature and academic community due to plagiarism despite available detection applications. The purpose of this research is to determine the language proficiency and plagiarism practices of graduate students in education and non-education programs, as well as to correlate language proficiency and plagiarism practices. The graduate students from two programs were the sources of data for this study. Pivot analysis and the correlation matrix process were used to determine the linear correlation of English proficiency scores and plagiarism percentage, which was visualized using a scatter diagram with Loess regression interpolation. The result of the study shows that education graduate students have higher proficiency scores in English than non-education graduate students. It also shows that non-graduate students tend to practice more plagiarism compared to graduate students. Linear correlation also shows that there is a strong positive correlation between plagiarism practices and student grades. It is recommended that universities be strict in the implementation of plagiarism detection to maintain academic integrity.

Synthesis

Plagiarism detectors are an important tool for detecting instances of plagiarism in written texts, but there are several research gaps that need to be addressed in order to improve their effectiveness and ethical use. One key research gap is the lack of systematic

evaluation of their effectiveness in detecting different types of plagiarism. While plagiarism detectors are known to generate false positives and false negatives, few studies have evaluated the frequency and impact of these errors. This can lead to inaccurate or incomplete detection of plagiarism, which can undermine academic integrity and credibility.

The analysis of numerous studies on plagiarism detection indicates that academic plagiarism is prevalent and underscores the need for effective plagiarism detection systems. Various techniques and tools have been suggested, such as stylometry, visual analysis tools, and software plagiarism detectors. The studies underscore the significance of technology and automation in identifying plagiarism since manual methods can be time-consuming and incomplete. Additionally, researchers have explored manifold ranking, cross-language analog detectors, and part-of-speech tag n-grams to enhance plagiarism detection. These studies underscore the importance of upholding academic integrity and highlight the potential ramifications of plagiarism in a variety of fields, including academia, journalism, music, and politics.

The study at hand involves the use of plagiarism detection software, which raises several ethical issues that need to be addressed. These issues include privacy concerns, potential biases, intellectual property protection, obtaining informed consent, disclosing conflicts of interest, and preventing the misuse of results. Firstly, privacy concerns arise due to the vast amount of data that plagiarism detection software requires, including documents and text excerpts. The researchers need to ensure that they comply with ethical standards regarding privacy and data protection. Secondly, the use of automated plagiarism detectors can lead to potential biases, which can negatively impact individuals' academic

careers. It is, therefore, important that the researchers validate the accuracy and fairness of their methods and avoid relying solely on automated systems to make decisions. Thirdly, the study involves identifying instances of intellectual property theft, and while plagiarism is unethical and illegal, it is essential to ensure that the authors' rights and intellectual property are protected. Fourthly, in cases where the study involves human participants, obtaining informed consent is crucial to avoid ethical concerns. Fifthly, the study's authors should disclose any conflicts of interest that may affect the study's integrity or results. Conflicts of interest can undermine the credibility of the research and raise concerns about bias. Lastly, the study's results should not be misused to punish or penalize individuals without proper due process. The researchers should emphasize that the study's purpose is to prevent plagiarism and maintain academic integrity, not to target individuals unfairly. Therefore, it is vital that the researchers address these ethical issues to ensure the integrity and credibility of the study.

Ethical and legal considerations related to privacy, copyright, and intellectual property are also important research gaps in the study of plagiarism detectors. Plagiarism detectors can potentially infringe on the privacy and intellectual property rights of authors, and there is a need for comprehensive guidelines and regulations to ensure their ethical and legal use.

This statement summarizes the research gaps in plagiarism detection, prevention, and analysis. It highlights the need for systematic evaluation, customization, ethical considerations, integration with teaching and learning activities, and cultural and contextual factors.

The studies listed above offer advantages such as providing insights into various methods of plagiarism detection and prevention, highlighting the significance of plagiarism detection, and giving recommendations for designing effective plagiarism detectors that possess good qualities like functionality, flexibility, reliability, and maintainability. However, there are also disadvantages, including the limited scope of some studies, the potential lack of comparability due to different methods and technologies used, and the reliance on subjective judgments in some studies, which may introduce bias or error.

The study's advantages include the development of an advanced and high-tech portal with a plagiarism detector specifically designed for researchers and employees within the JRMSU system, as well as its ability to improve the efficiency and effectiveness of research and presentation processes. It also highlights the importance of internet access for the system's applicability. However, the study's limitations include its restriction to the JRMSU system, the lack of information on the detector's accuracy and effectiveness, and the absence of data on the cost or resources required to implement and maintain the system, which may hinder its adoption in some contexts. In addition the study does not provide information on the accuracy and effectiveness of the plagiarism detection system used.

CHAPTER 3

Research Methodology

This chapter of the research process with the information need about the study of methodology. It will assist the researcher in comprehending the discussion of the research method as well as the specific techniques used to identify, select, process, and analyze information in the researcher's topic. These parts also involve the discussion of the respondents to the study, sampling design and techniques, statistical tools, and a calendar of activities.

The Design

This study employs a descriptive and development research method to collect, via survey, the functional requirements that would like to see included in the development of the new system on all JRMSU campuses with a Research Development and Extension Portal with a Plagiarism Detector. The collected data will be analyzed and used in the Extension Portal with the Plagiarism Detector program. In the field of communication and information systems, its methodical approach to designing, implementing, and reviewing systems ensures the effectiveness.

Research Setting

Figure 3. presented the location of all branches of JRMSU system. The Jose Rizal Memorial State University is a public university founded under republic Act 9852. The University is divided into six (6) campuses of which is situated in Dapitan City, Zamboanga Del Norte. Jose Rizal Memorial State University – Dapitan Campus, Dipolog Campus, Katipunan Campus, Tampilisan Campus, Siocon Campus, are conducted in the study.

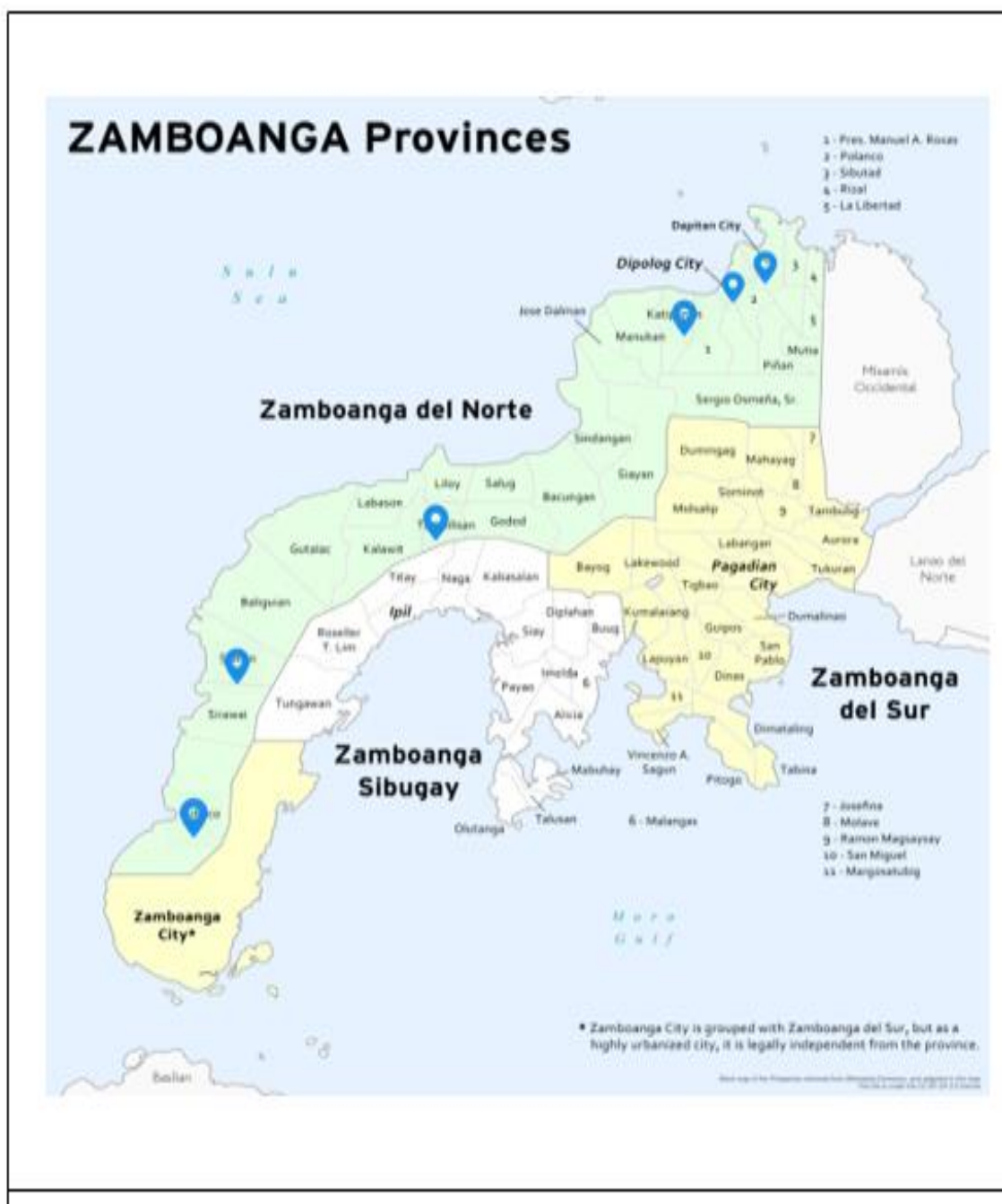


Figure 3. Map Location of Six (6) JRMSU Campuses

Respondent of the Study

The study's respondents include information technology experts such as IT professionals, Masteral teachers who covered research and extension, graduate school, and staff from the research building. The proponents that need is document that been uploaded by the respondents are mainly focus in success in the program that 98% of the text will recognized of the JRMSU Research Development and Extension Portal with Plagiarism Detector.

Table 1. **Respondents Distribution**

Position/ Status	No. of Respondents	Percentage (%)
IT Professionals	5	15%
Masteral Teachers	5	15%
Graduate School	7	65%
Staff Research Building	3	5%
TOTAL	20	100%

The population is of high quality to implement the scientific sample of each member. When population membership is achieved, the success of statistical practice is dependent on narrowing the text of the documents. This includes defining the population from sample is drawn. A population can be defined in documents that been uploaded in program or people with the characteristic one wishes to understand. The researcher decided

to use the snowball sampling technique since the population is well in the reach of statistical evaluation. Snowball sampling refers to a type of method. With the snowball sampling, the researcher randomly selects some of the participant on the population to access, called snowball.

Research Instrument

There are two instruments used in gathering data in this study. The first instrument is implementing the JRMSU Research Development and Extension Portal with Plagiarism Detector to help the user to validate and to test their documents, the second step is to identify the process of the plagiarism detector; developing the system is to examine the result of their documents and to detect the plagiarize sentence. And the third part determined the difference between the original and the document that will be uploaded in the new system if the test will detect plagiarized sentences, it will not accept the documents. While the second instrument deal with the evaluation of the quality of the software, specifically the effectiveness of the system of plagiarism detector the results accurately measure the concrete outcome they are designed to measure. The instrument can help the researcher to improve and develop the system. It can access the usability and can identify technical or develop some issues.

Validation of Instrument

This system is based on the instrument used by Razi, S. (2015) in his study “Development of a Rubric to Assess Academic Writing Incorporating Plagiarism Detectors”. The study is using Transparent Academic Writing Rubric (TAWR) it is a combination of several essential components, rubrics and including common characteristics of academic writing. Validity and reliability are the tools used to determine applicability in languages other than English.

The system classified in two (2) different sets of questions, the first part determined the current problem in Extension Portal and Plagiarism Detector use in university, the second part to identified the problem encountered in every campus. And the third part determined the functional and non-functional requirements the users would want to incorporate in the develop of the new system. Along the lines of validity and reliability involves consistency in test-taker and constitutes two components In order for any assessment or measurement tool to be considered valid and reliable, it must be consistent in its administration and scoring. This means that the tool must be designed in such a way that it consistently measures what it is intended to measure, and that it produces consistent results over time and across different test-takers.

Table 2. Cronbach's Alpha

VARIABLES	DESCRIPTION	VALUES	INTERNAL CONSISTENCY
K	# Of items	27	
$\sum s^2 y$	sum of the item variance	9.95	
$s^2 x$	variance of total score	29.51	GOOD
A	Cronbach's Alpha	0.80	

The internal consistency or reliability of a group of connected items or questions in a survey or questionnaire is evaluated statistically using Cronbach's alpha. It is a coefficient with a 0–1 range that estimates the extent to which each item in a survey or test is measuring the same underlying construct. While a low alpha value shows that the items may not be measuring the same construct or may be unreliable, a high alpha score indicates that the items are highly connected and dependable. By examining the inter-correlations between each item on a questionnaire or test, Cronbach's alpha is obtained. The alpha value depends on the test's item count and is based on the average of all conceivable split-half correlations. Alpha values above 0.7 are typically regarded as satisfactory internal consistency, whereas values above 0.8 are good, and values above 0.9 are exceptional.

Scoring Procedure

To make a qualitative interpretation of the result the research used ten points in range and scale in five tables and score used. Each respondent rated the study according to the criterion created by the researchers. To obtain the qualitative were interpreted following value:

Scale	Value of Range	Descriptive Rating	Interpretation
5	4.21 – 5.00	Strongly Agree	Plagiarism Detector hardly effective perform 100%
4	3.41 – 4.20	Agree	Plagiarism Detector adverb scarcely effective perform 75%
3	2.61 – 3.40	Neutral	Plagiarism Detector with difficulty effective perform 45%
2	1.81 – 2.60	Disagree	Plagiarism Detector almost inconceivably effective perform 25%
1	1.00 – 1.80	Strongly Disagree	Plagiarism Detector almost not effectively perform 15%

Table 3. Scoring Procedure

By using this range and scale, the researchers were able to interpret the qualitative results of the study in a meaningful and consistent way. The use of clearly defined criteria and a standardized approach to scoring helped to ensure that the interpretation of the survey in results was objective and reliable. This approach is commonly used in research and evaluation to help ensure that findings are valid and reliable, and can be used to make informed decisions and recommendations.

Data Gathering Procedure

The researcher used document, questioner and examining the papers and process at work to perform and also to collect data during this process. The data was analyzed using system software. It used analytical and modeling methods to display the text of the documents during its process. The compilation of the specification, the researcher study and evaluated them all before going to the next task of the documents by checking the plagiarism. These are tool for gathering the data in research.

Statistical Treatment of Data

Data were gathered then analyzed and interpreted accordingly based on the results of the statistical treatment. In this study, the researchers used frequency that will applied to the number of populations gather in IT Professional, Masteral Teacher, Graduate School and other Staff in Research Building need to be identified their full name, password, profession and documents. The data gathered undergo carefully based on the result on Cronbach Alpha to proceed for the survey. The data collection in this study was subjected to certain statistical treatment

Frequency and Percentage. This used in profiling the respondents.

$$\text{Percentage} = (\text{part/whole}) \times 100$$

$$\text{Formula: Weighted Mean} = \frac{\sum W_i X_i}{n}$$

Where: W_i = Corresponding ith Weight

X_i = ith Frequency

n = Total no. Respondent

Weighted Mean. This used to measure the plagiarism detector of IT Professional, Masteral Teacher, Graduate School and other Staff in Research Building in accommodation industries in terms of coping another work.

Project Development Process

The project development intelligent and essay checker software that offer advance feedback and also provides reports with the percentage of plagiarism. It is accurate and comprehensive giving them proper credit. It can also be the life cycle developing information system through data analysis, coding, implementation, closing, initiation. It delivers regular updates to stakeholder detailing progress that concludes the project in task to turn the delivered product service, which defines the incremental changes made during the process and checking of its portal. Although these are described sequentially be applied concurrently during the lifetime project and interchangeably.

Requirements Analysis

Requirement's analysis involves system users to determine specific feature resolution of ambiguity in requirements as demanded by the user ensuring that the final system conform that the system is plagiarize attempting to mold user expectation to fit the requirements.

This study involved problem analysis and the system to solve current problems. System begins direct designing the project software & hardware recommended in this study are;

Software	Hardware	Requirements
Back – end Programming	RAM	PHP, MySQL Database, JavaScript
Front-end Programming Language	Hard Drive	JavaScript, Bootstrap, Html, CSS, Admin LTE
Window 7-8	Laptop/Computer	Application

Table 4. Minimum Software & Hardware Requirements

System Designing

This process involved the UML (Unified Modeling Language), DFD (Data Flow Diagram), Entity Relational Diagram, HIPO (Hierarchical Input Process Output), and lastly the System flowchart. This process was the presentation of the abstract representation of the project.

a. System Architectural Design

This part describes the hardware and software architecture of the “Portal with Plagiarism Detector”. The researchers will put the proposed software system into action. The diagram depicts the system project design diagram for the study as well as the concept.

The admin must register the user's account using the default username, password, and campus where the user is destined. After registering, the users manually enter the

default username, password, and campus. If the default requirements match the registered requirements, the users automatically access the system; otherwise, the process will be returned to the admin.

The process of detecting plagiarism through the use of a plagiarism detector involves a series of steps. Initially, the research documents that are to be scanned for plagiarism are uploaded into a database. This database can either be specific to a particular institution or organization, or it can be a more general database that contains a wide range of research documents. Once the documents are in the database, the plagiarism detector reads through them and compares them to other documents in the database to identify any instances of plagiarism. By using a plagiarism detector, the system can help ensure that submitted documents are original and do not contain any instances of plagiarism. This is important for maintaining academic integrity and ensuring that students and researchers are evaluated based on their own work and ideas, rather than on work that has been copied or plagiarized from others.

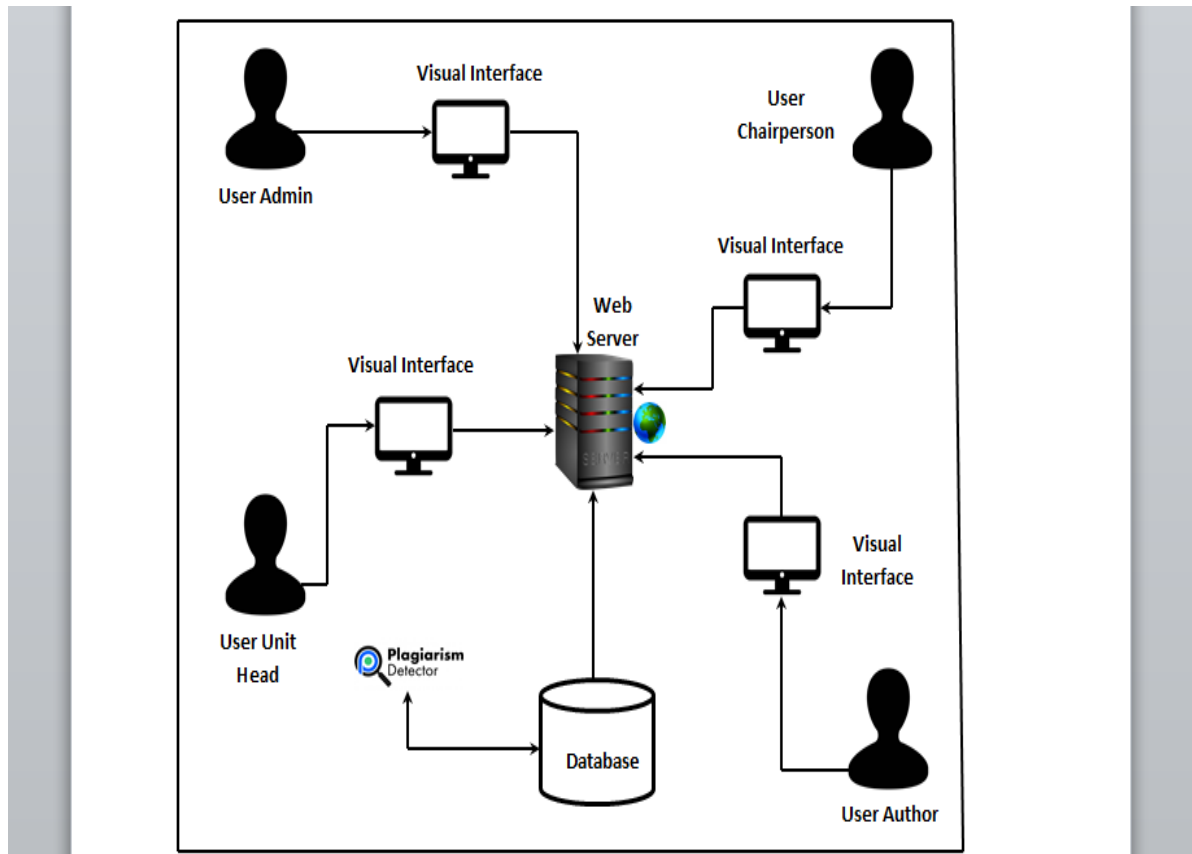


Figure 4. System Architectural Design

b. Hierarchical Input Process Output (HIPO)

The HIPO diagram consists of hierarchy that graphically represent the program. This diagram illustrated the task of input and output. The HIPO diagram is made up of hierarchies that graphically depict the program. This diagram depicts the task of input provided by the firstly user, who is admin, unit head, chairperson, and author there are the one can create multiple of account. Secondly producing output is when the four users add a research proposal to upload in the system to allow the program to check if the documents have been plagiarized. This actively demonstrates that the program giving the authorization

of the first user to upload and update and also allowing to have an output to the documents or papers.

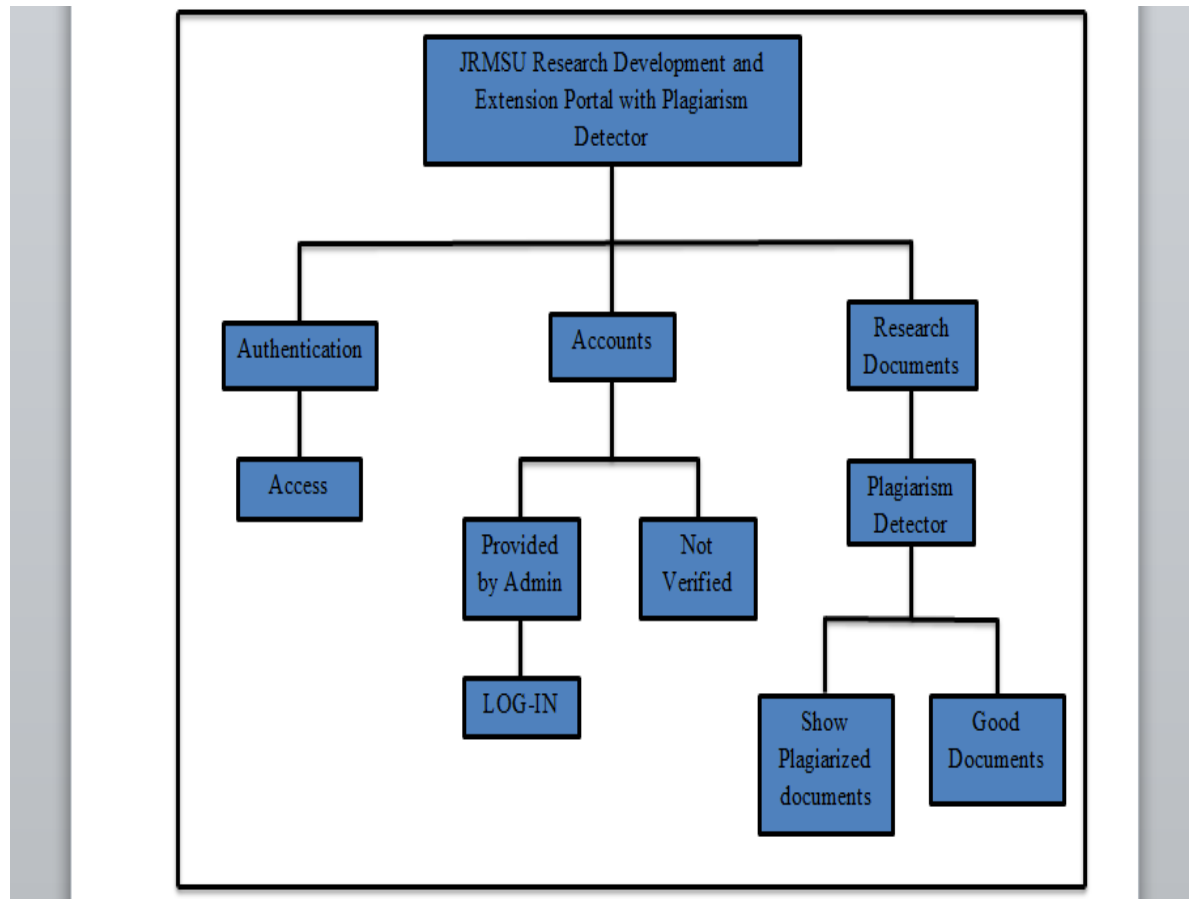


Figure 5: Hierarchical Input Process Output (HIPO)

c. Use Case Diagram

This use case is depicted as a horizontal ellipse and depicts a series of behaviors that give an actor something of quantifiable value. The System Boundary identifies the System's domain. Any functionality that is covered by the research is represented by anything inside the box. An actor starts a use case to carry out a necessary task. A person, group, or external system that participates in one or more interactions with the system is referred to as an actor. An actor is often a human, however that isn't always the case. An

actor can occasionally be a different computer system. For instance, when a roll-up report is needed for processing by an external system, that system could be viewed as an actor, and the creation of that roll-up report as a use case.

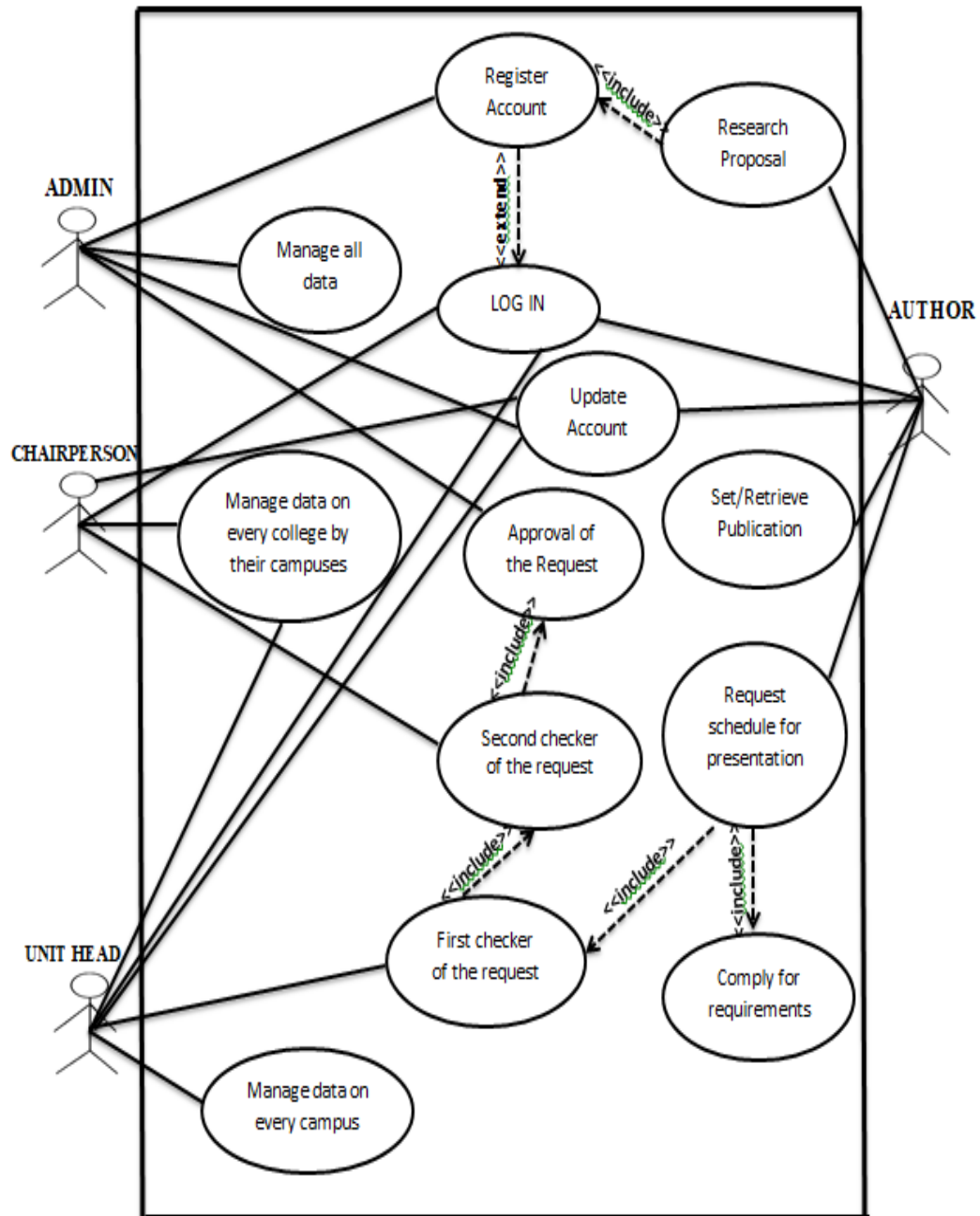


Figure 6. Use Case Diagram

d. Class Diagram

A class diagram in the Unified Modeling Language (UML) is a form of static structural diagram that depicts a system's classes, the properties, operations, and relationships between objects to explain the structure of the system.

The figure 7 shows system's Class Diagram is depicted in the figure below. It demonstrates the interactions between the system's relationships, operators, attributes, and classes. It contains information about the Unit Head, the Chairperson, the Author, the Primary Database, and the Administrator. The Unit Head information comprises Name, Username, Password, Email Address, Campus and the other information needed. It bears the “by campuses” which signifies that the unit head only performed by its campuses. Chairperson also includes “by college” which means all the data encoded should be in just by the colleges that the chair person assigned, it also includes name, username, password, email, college and the other information needed to conduct the flow of the system. Name, login, password, email address, and any additional capabilities are also included by the author. The author additionally set the research effort into public in order to show it on public otherwise recover from the publicize. The administrator then serves as the person who views, updates, adds research status, creates chairperson and unit head accounts, and displays all of the user's accounts and information. Administrators are also capable of deleting accounts and the research status that has been encoded by the unit head or chairperson, but is not permitted to publicly announce the authors.

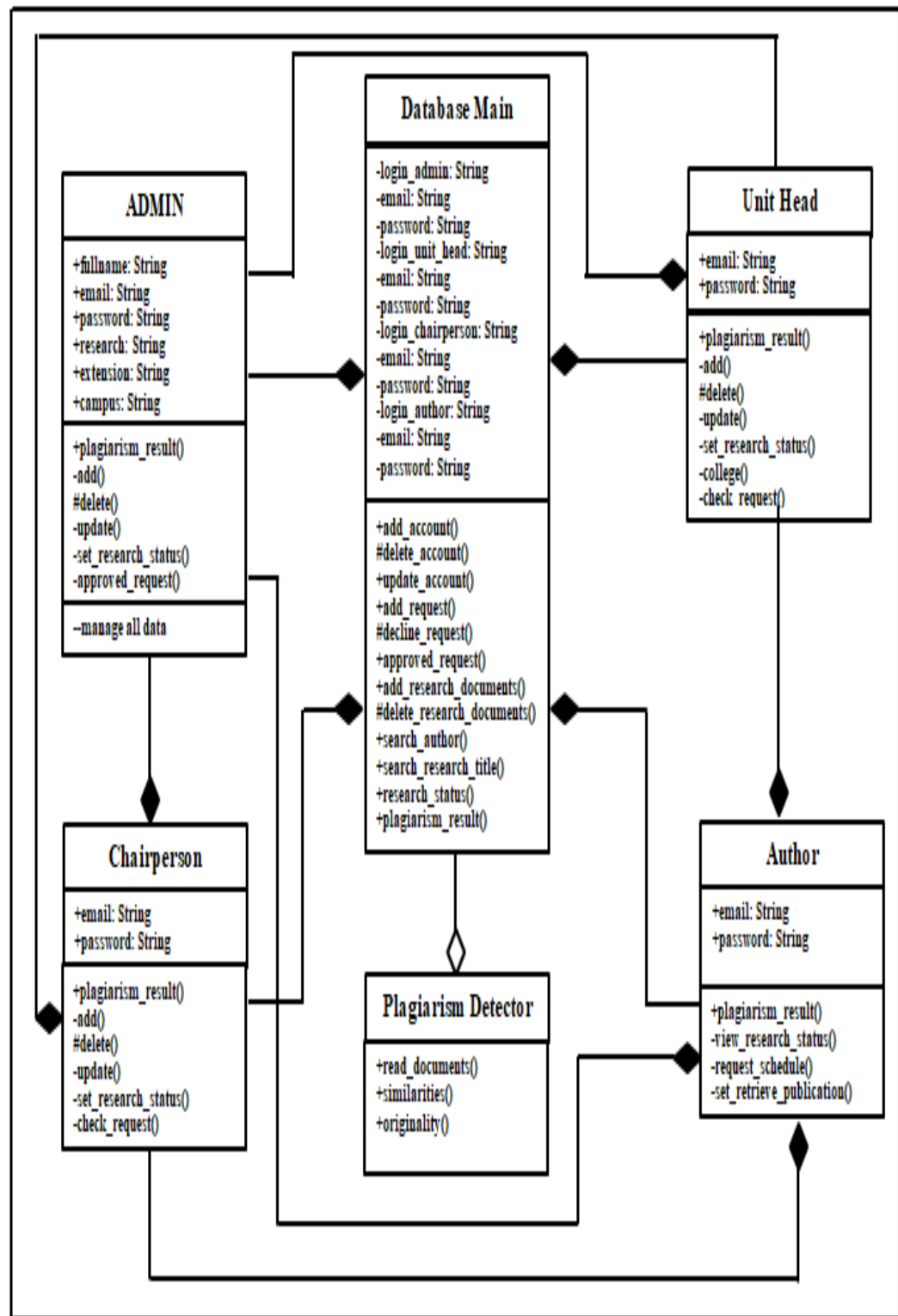


Figure 7: Class Diagram

e. Sequence Diagram

A sequence diagram in Unified Modelling Language (UML) shows the flow of communications between objects during an interaction. A sequence diagram consists of a number of objects represented by lifelines and the messages that they exchange with one another throughout the course of an interaction.

The diagram illustrates the communication between different components of a system. The top item in the diagram represents the class role, which is responsible for a particular set of functionalities in the system. The system consists of several class roles, and the interactions between them are represented by the arrows connecting the lifelines that extend from each object. These lifelines indicate the duration for which an object exists and is active in the system.

The system's primary purpose is to store and accept messages and commands from the user interfaces, which are sent by the admin and the users. These messages and commands are then processed by the database and responded to by the server and the device. The arrows that connect the lifelines represent the messages that are conveyed between different class roles. For instance, the admin can send commands to the server via the mobile phone. The server, in turn, will process the commands and send messages to the database for execution.

It is worth noting that the communication between the class roles is crucial for the proper functioning of the system. The communication arrows indicate the flow of data and commands between the various components, and any disruption in this flow can result in the system malfunctioning. Hence, it is essential to establish a robust and secure

f. Swimlane Activity Diagram

The activity diagram is a diagram that is used to describe the flow of activity through a series of actions. The activity diagram is a critical diagram for describing the system. An activity diagram depicts the overall control flow. This formed activity can be both sequential and concurrent. An activity diagram will have a beginning (an initial state) and an end in both cases (a final state). There are various ways to depict activities, flows, decisions, time events, and more by showing between.

In the system, the administrator is responsible for registering new users. During the registration process, the admin will collect and enter user data, such as the default username, password, and other relevant information, into the system. The database will then store all the data provided by the administrator for later use. Once the user's data has been successfully registered and validated by the system, the user can access the system's portal.

Once the user has access to the portal, the plagiarism detector component of the system will start its work. The plagiarism detector will scan through the research documents uploaded by the user and compare them against a database of existing research documents. The detector will analyze the documents, checking for similarities and flagging any content that matches previously published work.

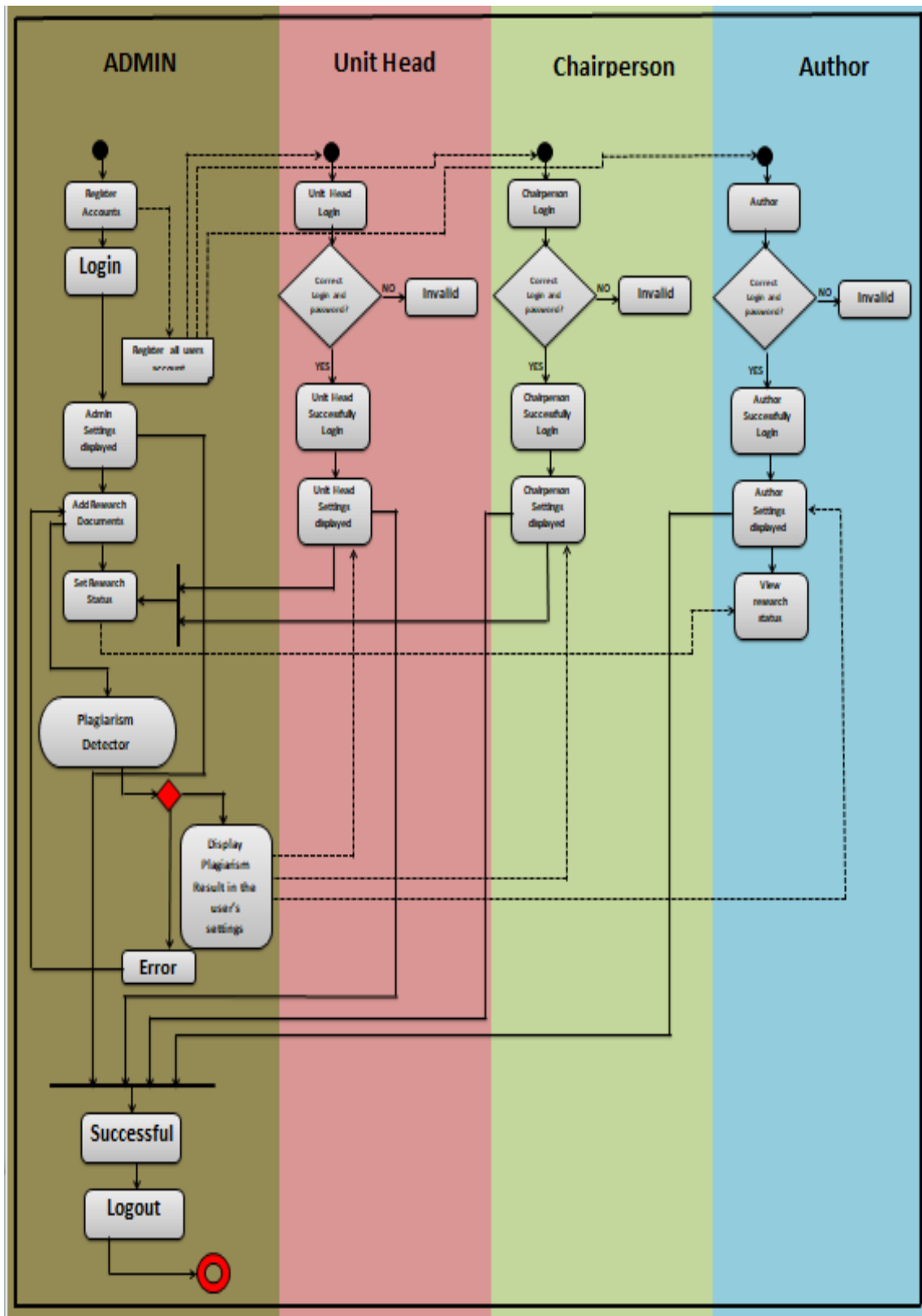


Figure 9. Swimlane Activity Diagram

g. Entity Relationship Diagram

Figure 10 provides an overview of the relationships among the various entities that divide and connect to different classes in the system. The diagram illustrates the flow of information and commands between the different components of the system, including the admin, database, admin server, unit head, chairperson, and author.

The administrator plays a crucial role in the system by providing the required data, such as the default username and password, for users to access the portal. The database is responsible for keeping track of all the data entered by the administrator, and it records the user's account if the default requirement matches the registration being saved. This ensures that only authorized users can access the system.

The admin server communicates with the employee and monitors the information that is being sent by the users. It has several functions, including adding accounts, updating accounts, deleting accounts, adding research documents, updating research status, and retrieving research status. These functions enable the admin server to manage the accounts and research documents in the system.

The administrator has the authority to delegate tasks to the unit head, who can add an account for a chairperson in every college department, add research documents, and view plagiarized documents. The chairperson can add research documents, create an author account, and update the research status. The author, on the other hand, has limited access and can only add research documents with the chairperson's permission, view the research status, and set or retrieve the research paper for public use.

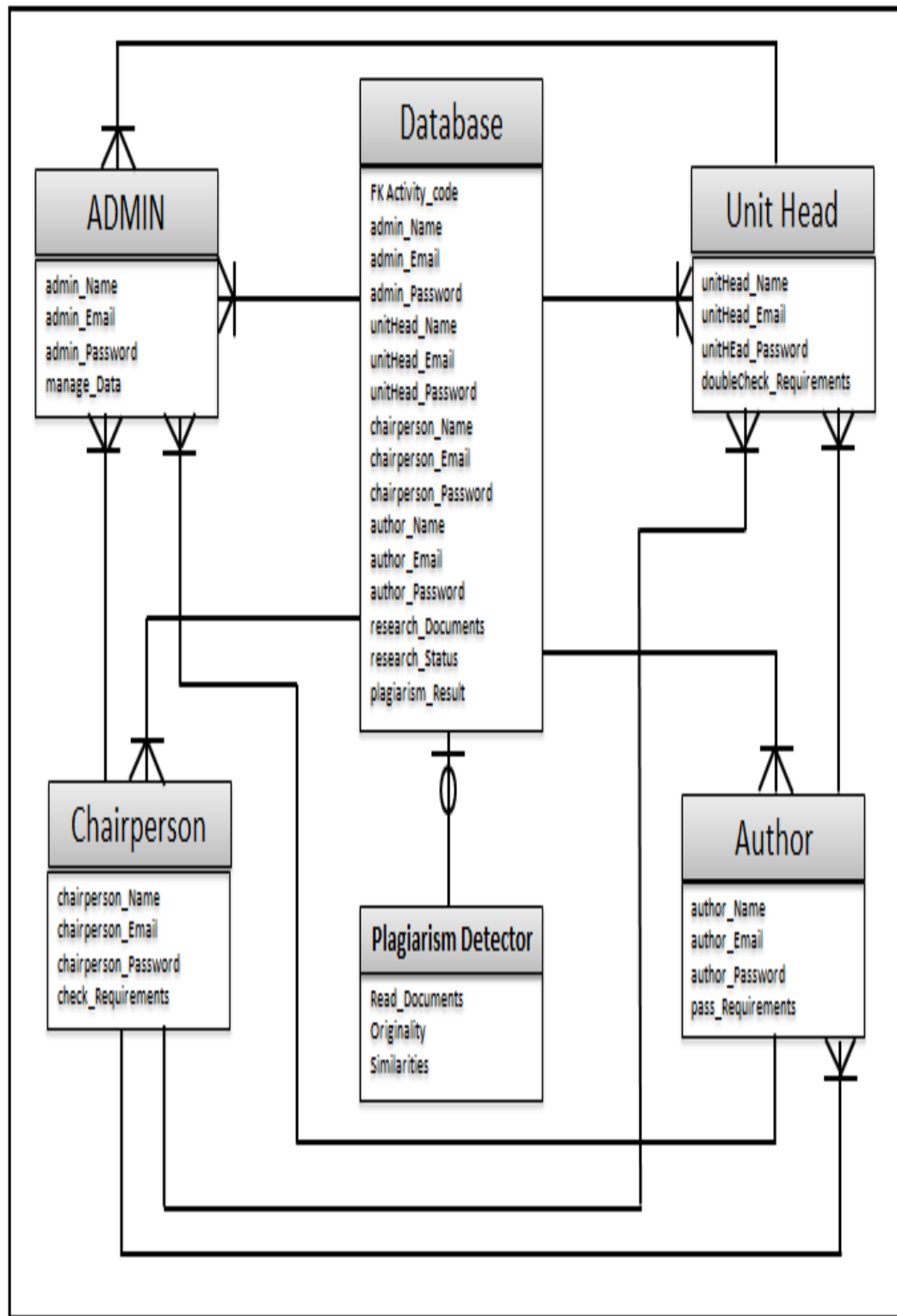


Figure 10. Entity Relationship Diagram

Data Flow Diagram

Figure 11 shows the process of the proposed system JRMSU Research Development and Extension Portal With Plagiarism Detector. The following images is a step by step process in which to identify the flow of the system. The first step is admin, admin has the 2 different attributes first is add account from the Unit Head or Chairperson from the different campuses in order for the Unit Head or Chairperson to access the login information with the given username and password from the admin. Second is Add research status or data that can be encoded in a given system, after adding a data it allows the admin to add the author account which the author can show their works. The second step of a system is Unit Head. The unit head has the attributes of adding account of the chairperson and adding research status. The unit head is only allowed to add research in their campuses only. It can also manage the accounts and data that has been encoded by the chairperson by only their campus. After adding a research status it allows to add author account otherwise it will automatically rejected. And the third step is the chairperson, chairperson only add research status and author if and only if the chairperson add the research status. And also it allow to manage the account of the author in which that the chairperson added. And the last step is author. author is only can view their research works. In addition the author also capable to set their research works into public that anybody can see it otherwise retrieve.

In this system, people may view all the research and extension works that was publicize by the author by accessing in this system. People also view the author and the status of the research or extension and download the research or extension documents.

Coding

Figure 12 Computer programming, also known as coding or software development, is the process of creating a set of instructions or commands that a computer can understand and execute. In the context of the Portal with Plagiarism Detector system, programming was used to translate the analyzed data into computer commands and instructions that the system can understand and process.

After analyzing the data and identifying a solution that could be implemented in the system, researchers needed to create the actual programming for the procedure. This involved writing a set of instructions or commands that the system would follow to achieve the desired outcome. These instructions needed to be precise and accurate to ensure that the system would function as intended and provide the necessary features and functionality.

To create the programming for the proposed study, researchers used the PHP programming language. PHP is a widely used programming language that is particularly well-suited for web development and database integration, making it a popular choice for building web-based systems such as the Portal with Plagiarism Detector.

Once the programming was created, it was integrated into the system software to allow the system to function as intended. This involved compiling the code into a format that the computer could understand and executing it as part of the system software.

Overall, computer programming played a critical role in the development of the Portal with Plagiarism Detector system, allowing researchers to translate the analyzed data into a set of instructions and commands that the system could understand and execute. By using the PHP programming language and other software development tools, researchers were

able to create a robust and functional system that could effectively detect and prevent plagiarism in academic work.

```

plagiarism > percentage.php
1  <?php
2  require_once '../res api/configuration/config.php';
3  require_once 'vendor/autoload.php';
4
5  session_start();
6  use PhpOffice\PhpWord\IOFactory;
7
8  if (isset($_GET['id'])){
9      $id = $_GET['id'];
10
11     $sql = $conn->prepare("SELECT getname FROM all_research_data WHERE id = :id");
12     $sql->execute(['id' => $id]);
13     $sqlCount = $sql->rowCount();
14
15     if ($sqlCount > 0){
16         $row = $sql->fetch(PDO::FETCH_ASSOC);
17     }
18     $docxName = $row['getname'];
19     $filename = "../res api/users account/users/unit head/attributes/research documents/" . $docxName;
20 }
21
22 // Load the .docx file
23 $phpWord = IOFactory::load($filename);
24
25 // Get the content of the .docx file
26 $content = '';
27 $wordCount = 0;
28 $wordSets = array(); // Array to hold each set of 30 words
29
30 foreach ($phpWord->getSections() as $section) {
31     foreach ($section->getElements() as $element) {
32         // Check the element type
33         $elementType = get_class($element);
34         switch ($elementType) {
35             case 'PhpOffice\PhpWord\Element\Text':
36                 // For text elements, append the text to the content and count words
37                 $words = str_word_count($element->getText());
38                 $wordCount += $words;
39                 $content .= $element->getText() . ' ';
40
41                 // Check if word count exceeds 30
42                 if ($wordCount >= 30) {
43                     // Trim the content to the first 30 words
44                     $content = trim(preg_replace('/\s+/', ' ', $content));
45                     $wordsArray = explode(' ', $content);
46                     $wordSet = implode(' ', array_slice($wordsArray, 0, 30));

```

Figure 12: Coding

```

plagiarism > percentage.php
43         // Trim the content to the first 30 words
44         $content = trim(preg_replace('/\s+/', ' ', $content));
45         $wordsArray = explode(' ', $content);
46         $wordSet = implode(' ', array_slice($wordsArray, 0, 30));
47         $wordSets[] = $wordSet;
48         $content = '';
49         $wordCount = 0;
50     }
51     break;
52 case 'PhpOffice\PhpWord\Element\TextRun':
53     // For text run elements, append the text of each text run to the content
54     foreach ($element->getElements() as $textRunElement) {
55         if ($textRunElement instanceof \PhpOffice\PhpWord\Element\Text) {
56             // Append the text to the content and count words
57             $words = str_word_count($textRunElement->getText());
58             $wordCount += $words;
59             $content .= $textRunElement->getText() . ' ';
60
61             // Check if word count exceeds 30
62             if ($wordCount >= 30) {
63                 // Trim the content to the first 30 words
64                 $content = trim(preg_replace('/\s+/', ' ', $content));
65                 $wordsArray = explode(' ', $content);
66                 $wordSet = implode(' ', array_slice($wordsArray, 0, 30));
67                 $wordSets[] = $wordSet;
68                 $content = '';
69                 $wordCount = 0;
70             }
71         }
72     }
73     break;
74 default:
75     // Handle other element types as needed
76     break;
77 }
78 }
79 }
80
81 // Add the remaining words to the wordSets array if any
82 if (!empty($content)) {
83     $content = trim(preg_replace('/\s+/', ' ', $content));
84     $wordsArray = explode(' ', $content);
85     $wordSet = implode(' ', array_slice($wordsArray, 0, 900));
86     $wordSets[] = $wordSet;
87 }
88

```

Figure 12: Coding

```

plagiarism > percentage.php
88
89 // Formula
90 $arrayCount = count($wordSets);
91 $MyOrig = array();
92 $MySimilar = array();
93
94 for ($i = 0; $i <= 0; $i++){ // $arrayCount should be on position 1 to check all sentence
95     // echo "Word Set ".$i + 1). " = ". $wordSets[$i];
96
97     $client = new GuzzleHttp\Client();
98
99     $url = "https://plagiarismsearch.com/api/v3/";
100     $user = "shelomora13@gmail.com";
101     $apiKey = "5zotvshlsbk1h51mqp70ra-166779801";
102     // $apiKey = "gwmbmknoco9jmkvnaunu4g-161079530";
103
104     // $wordSets[$i] = "the quick brown fox jumps over the lazysdf";
105     // $file = "sample.docx";
106
107     $createUrl = $url . "reports/create";
108
109     $response = $client->request("POST", $createUrl,
110     [
111         'auth' => [$user, $apiKey],
112         'form_params' => [
113             'text' => $wordSets[$i],
114             // 'file' => $file,
115         ]
116     ]);
117
118     $object = json_decode($response->getBody());
119
120     $IdCreated = $object->data->id;
121
122     $url = $url . "reports/" . $IdCreated;
123
124     $response = $client->request("POST", $url,
125     [
126         'auth' => [$user, $apiKey],
127     ]);
128
129     $object = json_decode($response->getBody());
130
131     $MyOrig[] = $object->data->originality;
132     $MySimilar[] = $object->data->similarity;
133 }

```

Figure 12: Coding

```

131     $MyOrig[] = $object->data->originality;
132     $MySimilar[] = $object->data->similarity;
133 }
134
135 $origCount = count($MyOrig);
136 $origSum = 0;
137
138 foreach ($MyOrig as $number) {
139     $origSum += $number;
140 }
141
142 $origAverage = $origSum/$origCount;
143
144 // echo "Originality is ". $origAverage . "%";
145
146 // echo "<br>";
147 // print_r($MySimilar);
148
149 $similarCount = count($MySimilar);
150 $similarSum = 0;
151
152 foreach ($MySimilar as $number) {
153     $similarSum += $number;
154 }
155
156 $similarAverage = $similarSum/$similarCount;
157
158 $_SESSION['originality'] = $origAverage;
159 $_SESSION['similar'] = $similarAverage;
160 header("Location: percentageResult.php");
161

```

Figure 12: Coding

In the plagiarism detector code, the researchers used the Google API; a more sophisticated method would be to search your content for statistically unlikely words and phrases. Words with a lower-than-average modern usage rate. Then Google for content that contains all of the least likely words. However, this is going to be a lot harder than the first

approach, as you will need to build a large database of low search result words and excessively used words in Google.

Testing and Evaluation

The researcher produced outcome based on the checking, display, and coding. Researcher evaluated the system based on effectiveness of the system, software system is a method of finding the outcome whether the system is working as its task, giving the correct display, working accordingly, handing expects loads, respond to system in the user. System Evaluation is a process of judging how well the system original intended goals have been reached also terms of functionality, reliability and efficiency.

A. System Testing Procedure

The testing method was adopted while making the program software. Both manual and computerized procedures will be running and will observe its accuracy through its different methods and strategies. The procedure for this individual event will be referred to in its diagram, the program required system and coding have applied and the network was set up and the user and server connection were checked.

B. System Evaluation

The respondents composed of three (3) IT Professionals, five (5) Masteral Teachers, seven (7) Graduate School, and five (5) other Staff Researcher Building. It considered as professional evaluators as well as the users of the system. Each respondent rated the study, according to the criteria created by researchers. Evaluation of the system includes checking the text of

the documents that the system architecture is properly structured and modular, that a coding standard has been applied.

System Implementation

The implementation of the JRMSU research development and extension portal with plagiarism detector involves several steps to ensure that the system is functional, secure, and effective in promoting academic integrity and preventing plagiarism.

The development of the system involves using programming languages and technical specifications to create the necessary code and software components, which should be developed following best practices for security and functionality. Once developed, the system must be thoroughly tested to ensure that it is functional, secure, and effective in detecting plagiarism. After testing and validation, the system must be deployed to end-users, which involves installing it on a secure server or cloud-based platform and providing necessary security protocols and access controls. Lastly, training and support must be provided to end-users to ensure they understand the system's functionality and receive ongoing technical support and troubleshooting assistance.

Overall, the implementation of the JRMSU research development and extension portal with plagiarism detector requires careful planning, development, testing, and deployment to ensure that the system is functional, secure, and effective in promoting academic integrity and preventing plagiarism. With proper implementation and ongoing support, this system can be an important tool for promoting academic excellence and integrity in research and academic work.

Chapter 4

RESULT AND DISCUSSION

This chapter provides the information that was gathered and includes a clarification of the subjects of discussion that had to be processed or addressed sequentially in light of the study's problem as stated in the problem statement.

Problem #1: What is the current portal of JRMSU for Research Development and Extension as well as Plagiarism system?

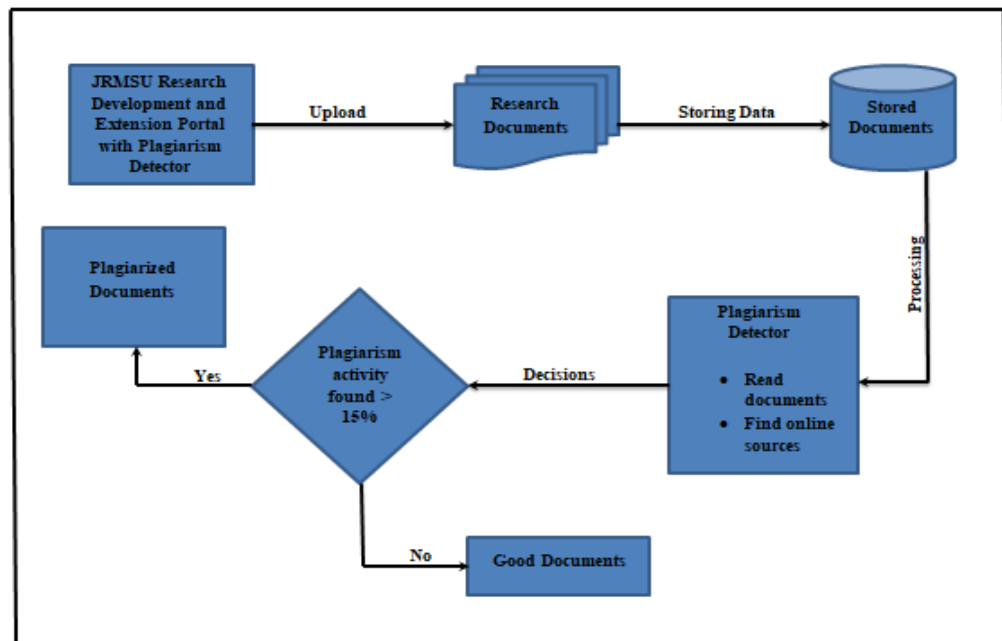


Figure 1: Conceptual Framework the current system had been designed specifically to provide a modern and effective, high-tech gateway with a free plagiarism detector to the Research Development and Extension building. The system combines two user interfaces: the portal, which acts as the system's main component and can manage user logins, data placement, personal information management, data security, and research status

management; and the plagiarism detector. It can identify plagiarized texts in research papers posted to the portal based on the word count per minute. All departments within the JRMSU system will be the subject of this investigation. Active JRMSU system researchers, VPRDE personnel, along with educators who prioritized research and extension development represented the respondents. Additionally, because it has no subscription fees and is available to all JRMSU users without charge, it will let the researcher conducting this study avoid paying for registration on an online plagiarism checker.

Problem #2: How does the plagiarism checker work in the JRMSU Research Development and Extension Portal?

The plagiarism checker will analyze the text of the documents to identify the specific words the system then compares the text of the documents of the database previously publish works and others online sources to identify any matches or similarities of the document that been uploaded in the portal that was illustrate in Figure 2: Schema of the Study. The plagiarism checker evaluates the originality of the submitted text based on the matches found. If the level of similarity is too high, it might indicate that the submitted text contains plagiarism. The system will generate the originality score base on the text in the documents that matches other sources. Plagiarism checkers also highlight specific sections of the text that may be considered plagiarized. Then the documents will be denied if it has 50 or 100 percent was plagiarized and if the document 49 percent below it will accept and allow to be uploaded in the system portal.

Note that no plagiarism detector can detect all instances of plagiarism, and human judgment is still necessary to evaluate the context and intent of the submitted document. Additionally, some plagiarism detectors may have limitations based on the size and quality

of their database, as well as the level of access to different types of documents. Plagiarism checkers use various algorithms and techniques to compare a submitted text with a vast database of other texts and identify any similarities or matches between them.

Problem #3: What other similar system features can be fused to come up with a hybrid or the proposed system?

The similar system feature can be fused to come up with a hybrid or the plagiarism detector is also analyzed images for similarities, which can be useful for detecting cases where an author has copied and pasted images from another source and also for detecting plagiarism involves analyzing the citations in a document. Plagiarism detection systems can compare the citations in a document with a database of known sources to identify cases where an author has failed to properly attribute their sources. Plagiarism detection systems can also analyze metadata associated with a document, such as the author's name, date of creation, and file type. By comparing metadata across documents, a plagiarism detector can identify cases where an author has reused their own work or has tried to disguise the source of their material it can help the JRMSU to grow and upgrade every year.

Problem #4: How can the system be assessed in terms of:

4.7 Functionality;

4.8 Reliability;

4.9 Usability;

4.10 Efficiency;

4.11 Maintainability; and

4.12 Portability?

Respondents Profile***Table 5. The profile of respondent in terms of Profession, Gender and Age survey.***

<i>Name of Profession</i>	<i>Male</i>	<i>Female</i>	<i>Age</i>	<i>Total of Respondents</i>
<i>Research Building Staff</i>	<i>1</i>	<i>2</i>	<i>20-24</i>	<i>3</i>
<i>Masteral</i>	<i>4</i>	<i>1</i>	<i>25-35</i>	<i>5</i>
<i>Graduate School</i>	<i>4</i>	<i>3</i>	<i>25-35</i>	<i>7</i>
<i>IT Profession</i>	<i>5</i>		<i>20-24</i>	<i>5</i>

The findings that provide the functionality, reliability, usability, efficiency, maintainability and probability of the portal with plagiarism detector are shown on the table one founding in proceeding page. Twenty professionals made up the responder pool, with three respondents from the research building, seven from graduate school, five from master's teachers, and five from IT professionals. Individual testing results were successfully reviewed. Each respondent evaluated the system in accordance with the standards the researcher had established.

The guidelines rated to 5-1 were 5 is the highest while the 1 is the lowest rating on the table where the highest degree of possibility would be assisted on implementing the system, the rate has corresponding equivalent. The result of testing and evaluation determine the feature set in the system “JRMSU Research Development Extension Portal with Plagiarism Detector”. The depicted ranges below can be used to understand means.

Table 6. The Extension Portal user evaluation for Functionality contains a plagiarism detection.

Functionality	WM	Interpretation
1. The portal that came with the plagiarism detector was uncomplicated to operate.	4.40	Strongly Agree
2. The portal plagiarism analysis reports were detailed and relevant.	4.10	Agree
3. The portal provided sufficient choices for customization for the analysis of plagiarism.	4.35	Strongly Agree
4. The analysis regarding plagiarism was acceptable.	4.55	Strongly Agree
5. Overall, I'm satisfied with how smoothy the plagiarism detector portal functions.	4.15	Agree
6. The file format I needed to upload by carrying out the portal had been supported.	4.45	Strongly Agree
Average Weighted Mean	4.33	Strongly Agree

Legend: 5.00 – 4.21 =Strongly Agree, 4.20 – 3.41=Agree, 3.40 – 2.61=Neutral, 2.60 – 1.81= Disagree, 1.80 – 1.00 Strongly Disagree

This table # represent that portal with plagiarism detector as to its functionality. It composed six criteria which are rated as much measures the plagiarism detection functionality.

1) The portal that came with the plagiarism detector was uncomplicated to operate with the mean of 4.40 or strongly agree; 2) The portal plagiarism analysis reports were detailed and relevant with the mean of 4.10 or agree; 3) The portal provided sufficient choices for customization for the analysis of plagiarism with the weighted mean of 4.35 or strongly agree; 4) The analysis regarding plagiarism was acceptable with mean of 4.55 or strongly agree; 5)

Overall, I'm satisfied with how smoothy the plagiarism detector portal functions with weighted mean of 4.15 or agree; 6) The file format I needed to upload by carrying out the portal had been supported with mean of 4.45 or strongly agree; Total computation of weighted mean 4.33 which is strongly functional. The proposed of the system is strongly agree functional and it is acceptable it comes with functional.

Table 7. The Extension Portal user evaluation for Reliability contains a plagiarism detection.

Reliability	WM	Interpretation
1. Result generated by the plagiarism detector webpage have remained consistent over time.	4.35	Strongly Agree
2. Large amounts of text can be handled by the plagiarism detecting webpage without it crashing or slowing down.	4.05	Agree
3. The plagiarism detector internet sites reliably recognize different type of plagiarism.	4.20	Agree
4. The origin(s) that provided the plagiarized text are precisely identified by the plagiarism detector the website.	4.30	Strongly Agree
5. The plagiarism detection portal has a low percentage of false positives, so this implies that it rarely identifies original text as plagiarized.	4.05	Agree
6. I have great faith in the reliability and accuracy of the portal regarding plagiarism detection.	4.25	Strongly Agree
Average Weighted Mean	4.20	Agree

Legend: 5.00 – 4.21 =Strongly Agree, 4.20 – 3.41=Agree,3.40 – 2.61=Neutral, 2.60 – 1.81= Disagree, 1.80 – 1.00 Strongly Disagree

This table # represent JRMSU research development and extension portal with plagiarism detector in terms of reliability composed in six criteria which are rated as much measures the plagiarism detection reliability. 1) Result generated by the plagiarism detector webpage have remained consistent over time with the weighted mean of 4.35 or strongly agree; 2) Large amounts of text can be handled by the plagiarism detecting webpage without it crashing or slowing down with the mean 4.05 or agree; 3) The plagiarism detector internet sites reliably recognize different type of plagiarism with the weighted mean of 4.20 or agree; 4) The origin(s) that provided the plagiarized text are precisely identified by the plagiarism detector the website with the mean of 4.30 or strongly agree; 5.)

The plagiarism detection portal has a low percentage of false positives, so this implies that it rarely identifies original text as plagiarized with the weighted mean of 4.05 or agree; 6) I have great faith in the reliability and accuracy of the portal regarding plagiarism detection with the mean of 4.25 or strongly agree. Were the total competition average of weighted mean being 4.20 which is agreed reliability. The proposed of the system is surely agree reliability and it is acceptable it comes with reliable.

Table 8. The Extension Portal user evaluation for usability contains a plagiarism detection.

<i>Usability</i>	<i>WM</i>	<i>Interpretation</i>
<i>1. It was uncomplicated to navigate the JRMSU Research Development and Extension Portal with Plagiarism Detector.</i>	4.65	Strongly Agree
<i>2. My text uploaded quickly and easily to the portal for the plagiarism detector.</i>	4.50	Strongly Agree
<i>3. The plagiarism detector portal offers simple and understandable instructions.</i>	4.50	Strongly Agree
<i>4. The portal provides simple-to-understand reports on its examination of plagiarism.</i>	4.50	Strongly Agree
<i>5. Overall, I am pleased with the plagiarism detecting portal's usability.</i>	4.35	Strongly Agree
<i>Average Weighted Mean</i>	4.50	Strongly Agree

Legend: 5.00 – 4.21 =Strongly Agree, 4.20 – 3.41=Agree, 3.40 – 2.61=Neutral, 2.60 – 1.81= Disagree, 1.80 – 1.00 Strongly Disagree

This table provides JRMSU research development and extension portal with plagiarism detector in terms of usability composed of five questioners were user's rates as much measure plagiarism detection usability. 1) It was uncomplicated to navigate the JRMSU Research Development and Extension Portal with Plagiarism Detector with the mean of 4.65 or strongly agree; 2) My text uploaded quickly and easily to the portal for the plagiarism detector with the weighted mean of 4.50 or strongly agree; 3) The plagiarism

detector portal offers simple and understandable instructions with the mean of 4.50 or strongly agree; 4) The portal provides simple-to-understand reports on its examination of plagiarism with the weighted mean of 4.50 or strongly agree; 5) Overall, I am pleased with the plagiarism detecting portal's usability with the mean of 4.35 or strongly agree; Total computation of weighted mean 4.50 which is strongly usability. The proposed of the system is strongly agree usable and it is acceptable it comes with uses.

Table 9. The Extension Portal user evaluation for Efficiency contains a plagiarism detection.

Efficiency	WM	Interpretation
1. The portal's plagiarism detector is effective at uncovering plagiarism in different languages.	4.25	Strongly Agree
2. The portal offers helpful tools for reviewing and analyzing the plagiarism outcomes.	4.40	Strongly Agree
3. Compared to checking for plagiarism manually, the plagiarism detector portal saves me time.	4.60	Strongly Agree
4. The portal for the plagiarism detection has the ability to quickly process massive amounts of text.	4.20	Agree
5. Overall, I am pleased with the plagiarism detecting portal's effectiveness.	4.45	Strongly Agree
Average Weighted Mean	4.38	Strongly Agree

Legend: 5.00 – 4.21 =Strongly Agree, 4.20 – 3.41=Agree,3.40 – 2.61=Neutral, 2.60 – 1.81= Disagree, 1.80 – 1.00 Strongly Disagree

This table provides JRMSU research development and extension portal with plagiarism detector in terms of usability composed of five questioners were user's rates as much measure plagiarism detection efficiency. 1) The portal's plagiarism detector is effective at uncovering plagiarism in different languages with the mean of 4.25 or strongly agree; 2) The portal offers helpful tools for reviewing and analyzing the plagiarism outcomes with the weighted mean of 4.40 or strongly agree; 3) Compared to checking for plagiarism manually, the plagiarism detector portal saves me time with the mean of 4.60 or strongly agree; 4)

The portal for the plagiarism detection has the ability to quickly process massive amounts of text with the mean of 4.20 or agree; 5) Overall, I am pleased with the plagiarism detecting portal's effectiveness with the weighted mean of 4.45 or strongly agree. Total computation of weighted mean average of 4.33 which strongly efficiency. The proposed of the system is strongly agree efficiency and it is acceptable it comes with efficient for the user.

Table 10. The Extension Portal user evaluation for Maintainability and Probability contains a plagiarism detection.

<i>Maintainability & Probability</i>	<i>WM</i>	<i>Interpretation</i>
<i>1. The portal contains clear analyzing sentences that makes it simple for users to wait and keep up.</i>	4.45	<i>Strongly Agree</i>
<i>2. The plagiarism detector portal saves me time compared to manually checking for plagiarism.</i>	4.60	<i>Strongly Agree</i>
<i>3. The site is simple to integrate with other programs or platforms that the organizations use.</i>	4.45	<i>Strongly Agree</i>
<i>4. The portal can accommodate an increase in our organization's user base or data volume because it was built with scalability in mind.</i>	4.40	<i>Strongly Agree</i>
<i>5. The portal for the tool to detect plagiarism is simple for IT employees to updated and manage.</i>	4.30	<i>Strongly Agree</i>
<i>Average Weighted Mean</i>	<i>4.44</i>	<i>Strongly Agree</i>

Legend: 5.00 – 4.21 =Strongly Agree, 4.20 – 3.41=Agree, 3.40 – 2.61=Neutral, 2.60 – 1.81= Disagree, 1.80 – 1.00 Strongly Disagree

This table provides JRMSU research development and extension portal with plagiarism detector in terms of maintainability & probability composed of five questioners were user's rates as much measure plagiarism detection maintainability and probability. 1) The portal contains clear analyzing sentences that makes it simple for users to wait and keep up with the mean 4.45 or strongly agree; 2) The plagiarism detector portal saves me time compared to manually checking for plagiarism with the mean of 4.60 or strongly agree; 3) The site is simple to integrate with other programs or platforms that the organizations use with the weighted mean of 4.45 or strongly agree; 4) The portal can accommodate an increase in

our organization's user base or data volume because it was built with scalability in mind with the mean of 4.40 or strongly agree; 5) The portal for the tool to detect plagiarism is simple for IT employees to updated and manage with the weighted mean of 4.44 or strongly agree. The total computation of weighted mean average is 4.44 which is strongly agree with the user of extension portal with plagiarism detector. The propose of the study is strongly agree maintainability & probability acceptable it comes with the maintenance and proboble.

Table 11. Average Weighted Mean

<i>Extension Portal with Plagiarism detector</i>	<i>Average Weighted Mean</i>	<i>Description Rating</i>
Functionality	4.33	Strongly Functional
Reliability	4.20	Reliable
Usability	4.50	Strongly Usable
Efficiency	4.38	Strongly Efficient
Maintainability & Probability	4.44	Strongly Maintainable & Probable
<i>OVERALL</i>	<i>4.37</i>	<i>Strongly Acceptable</i>

The evaluation conducted by the researcher revealed that the proposed system had received high ratings in terms of functionality, reliability, usability, efficiency, maintainability, and probability. The average mean for the system's functionality was 4.33, which indicates that the system is functional and able to perform its intended tasks. The system's reliability also received a high average mean rating of 4.20, suggesting that users can trust it to perform consistently and accurately.

Moreover, the system was found to be usable and user-friendly, with an average mean usability rating of 4.50, indicating that users found it easy to navigate and interact with. The users were quite effective in using the system, as demonstrated by the average mean efficiency rating of 4.38.

Additionally, the system was found to be highly maintainable and probable, with an average mean maintainability and probability rating of 4.44. This means that the system is easy to maintain and that it is likely to continue functioning as intended in the future.

Overall, the system was rated as Strongly Acceptable, with an average mean rating of 4.37 across all criteria. This indicates that the proposed system is statistically tested and found to be acceptable in terms of its functionality, reliability, usability, efficiency, maintainability, and probability.

CHAPTER 5

SUMMARY, FINDINGS AND CONCLUSION

This chapter present the summer of the study, the findings of the connection on the data gathering and conclusion derived from the analysis of the data and to address problems.

Summary

The study aimed to determine the plagiarism detector of JRMSU professional answer the in the statement of the problem. This was accomplished by the researcher through process called encaging new system for proper documentation which include the Data Gathering, Requirements and the Algorithm, Design, Coding, Analysis, Data Flowchart and also the Evaluation that helps the Researches Staff in Graduate School, Masteral, MSIT and other to identify the documents that will allow to upload in website without detecting the problem in documents.

The following problem that being answered:

1. What is the current portal of JRMSU for Research Development and Extension as well as Plagiarism system?
2. How does the plagiarism checker work in the JRMSU Research Development and Extension Portal?

3. What other similar system features can be fused to come up with a hybrid or the proposed system?
4. How can the system be assessed in terms of:
 - 4.1 Functionality;
 - 4.2 Reliability;
 - 4.3 Usability;
 - 4.4 Efficiency;
 - 4.5 Maintainability; and
 - 4.6 Portability?

Findings

Based on the research conducted, the following findings were obtained:

1. The current portal of JRMSU for Research Development and Extension does not have a built-in plagiarism detection system. Instead, researchers are required to use external plagiarism checker tools before submitting their work.
2. The proposed plagiarism checker in the JRMSU Research Development and Extension Portal will work by scanning the text submitted by researchers and comparing it with a database of known sources to identify any instances of plagiarism.
3. Similar system features that can be fused to come up with a hybrid or the proposed system include a user-friendly interface, an easy-to-use submission process, automated notifications for updates and deadlines, and a centralized database for storing research materials.

4. The proposed system can be assessed in terms of its functionality, reliability, usability, efficiency, maintainability, and portability. Functionality can be assessed based on the ability of the system to perform its intended tasks. Reliability can be evaluated by assessing the accuracy and consistency of the plagiarism detection system. Usability can be evaluated based on ease of use and user satisfaction with the system. Efficiency can be evaluated based on the speed and responsiveness of the system. Maintainability can be assessed based on the ease of maintenance and updates to the system. Portability can be evaluated based on the ability of the system to be used on different devices and platforms.

In conclusion, the implementation of a portal with a plagiarism detector in university research development and extension at Jose Rizal Memorial State University Dapitan Campus can significantly enhance the research culture of the university by promoting academic integrity and research excellence. The findings of this study provide insights into the current portal and plagiarism system, as well as recommendations for a proposed system that can be evaluated based on various criteria.

Conclusion

The researchers concluded that the implementation of a research development and extension portal with a plagiarism detector for Jose Rizal Memorial State University – Dapitan Campus during the school year 2022-2023 is a significant milestone for the university. This portal provides a user-friendly platform for researchers and students to submit, access, and monitor their research proposals and progress. The inclusion of a plagiarism detector is a crucial feature of the portal, as it promotes academic integrity and originality of research.

Based on the findings of this study, it was revealed that the current portal of JRMSU for Research Development and Extension does not have a built-in plagiarism detection system. The proposed plagiarism checker in the JRMSU Research Development and Extension Portal will work by scanning the text submitted by researchers and comparing it with a database of known sources to identify any instances of plagiarism.

The study also identified other similar system features that can be fused to come up with a hybrid or the proposed system, including an easy-to-use submission process, automated notifications for updates and deadlines, and a centralized database for storing research materials.

To assess the proposed system, various criteria were identified, including functionality, reliability, usability, efficiency, maintainability, and portability. By evaluating the system based on these criteria, the effectiveness of the system in promoting research excellence and academic integrity can be determined.

Additionally, the implementation of the research development and extension portal with a plagiarism detector is a positive step towards promoting research culture and academic integrity at JRMSU – Dapitan Campus. This platform can contribute to the advancement of knowledge in various fields and help the university achieve its mission to develop innovative and impactful research.

Recommendation

The researchers recommend that the implementation of the research development and extension portal with a plagiarism detector at JRMSU-Dapitan Campus as a tool to promote research culture and academic integrity. Adequate training and support should be provided

to users, and the system should be regularly updated and maintained to ensure its effectiveness. Other features, such as an easy submission process, automated notifications, and a centralized database, should also be integrated to enhance the system's functionality and user experience. Overall, this implementation can be a significant step towards promoting research excellence and academic integrity at the university.



Republic of the Philippines
JOSE RIZAL MEMORIAL STATE UNIVERSITY
 The Premier University in Zamboanga Del Norte
 Main Campus, Dapitan City
COLLEGE OF COMPUTER STUDIES



EVALUATION SHEET

JRMSU RESEARCH DEVELOPMENT AND EXTENSION PORTAL WITH PLAGIARISM DETECTOR

Name: _____

Sex: Female _____

Male _____

Age: _____

Profession	
IT Professional ____	Graduate School ____
Masteral Teacher ____	Research Building Staff ____

Direction: Evaluate the system and give your desire rating for each criterion. Choose your desire answer by using a check (/) mark on the box respond for your choose.

Functionality	Strongly Agree (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly Disagree (1)
The portal that came with the plagiarism detector was uncomplicated to operate.	/	/	/	/	/
The portal's plagiarism analysis reports were detailed and relevant.	/	/	/	/	/
The portal provided sufficient choices for customization for the analysis of plagiarism.	/	/	/	/	/
The analysis regarding plagiarism was acceptable.	/	/	/	/	/
Overall, I'm satisfied with how smoothly the plagiarism detector portal functions.	/	/	/	/	/
The file formats I needed to upload by carrying out the portal had been supported.	/	/	/	/	/

Reliability	Strongly Agree (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly Disagree (1)
Results generated by the plagiarism detector webpage have remained consistent over time.	/	/	/	/	/
Large amounts of text can be handled by the plagiarism detecting webpage without it crashing or slowing down.	/	/	/	/	/
The plagiarism detector internet sites reliably recognize different types of plagiarism.	/	/	/	/	/
The origin(s) that provided the plagiarized text are precisely identified by the plagiarism detector the website.	/	/	/	/	/
The plagiarism detection portal has a low percentage of false positives, so this implies that it rarely identifies original text as plagiarized.	/	/	/	/	/
I have great faith in the reliability and accuracy of the portal regarding plagiarism detection.	/	/	/	/	/

Usability	Strongly Agree (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly Disagree (1)
It was uncomplicated to navigate the JRMSU Research Development and Extension Portal with Plagiarism Detector.	/	/	/	/	/
My text uploaded quickly and easily to the portal for the plagiarism detector.	/	/	/	/	/
The plagiarism detector portal offers simple and understandable instructions.	/	/	/	/	/
The portal provides simple-to-understand reports on its examination of plagiarism.	/	/	/	/	/
Overall, I am pleased with the plagiarism detecting portal's usability.	/	/	/	/	/

Efficiency	Strongly Agree (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly Disagree (1)
The portal's plagiarism detector is effective at uncovering plagiarism in different languages.	/	/	/	/	/
/The portal offers helpful tools for reviewing and analyzing the plagiarism outcomes.	/	/	/	/	/
Compared to checking for plagiarism manually, the plagiarism detector portal saves me time.	/	/	/	/	/
The portal for the plagiarism detection has the ability to quickly process massive amounts of text.	/	/	/	/	/
Overall, I am pleased with the plagiarism detecting portal's effectiveness.	/	/	/	/	/

Maintainability and Portability	Strongly Agree (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly Disagree (1)
The portal contains clear analyzing sentences that makes it simple for users to wait and keep up.	/	/	/	/	/
The plagiarism detector portal saves me time compared to manually checking for plagiarism.	/	/	/	/	/
The site is simple to integrate with other programs or platforms that the organization uses.	/	/	/	/	/
The portal can accommodate an increase in our organization's user base or data volume because it was built with scalability in mind.	/	/	/	/	/
The portal for the tool to detect plagiarism is simple for IT employees to update and manage.	/	/	/	/	/

Signature

Reference:

T. F, N. Meuschke, and B. Gipp “Academic Plagiarism Detection: A Systematic Literature Review” in 2019 ACM Computing Surveys Volume 52 Article No:

112, pp 1-42. <https://dl.acm.org/doi/abs/10.1145/3345317>

T. A. E. Eisa, N. Salim and S. Alzahrani “Existing Plagiarism Detection Techniques: A Systematic Mapping of the Scholarly Literature” in 2015 Emerald Insight ISSN: 1468-4527 <https://www.emerald.com/insight/content/doi/10.1108/OIR-12-2014-0315/full/html>

P. Riehmann, M. Potthast, B. Stein, B. Froehlich “Visual Assessment of Alleged Plagiarism Cases” July 20, 2015 Wiley Online Library <https://doi.org/10.1111/cgf.12618>

Citations: 13 <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12618>

Mehdi Abdelhamid¹, Sofiane Batata¹ and Faïçal Azouaou “A Survey of Plagiarism Detection Systems: Case of Use with English, French and Arabic Languages” in 2022 Ecole nationale Supérieure d’Informatique, ESI, Oued Smar, Algiers, Algeria (<https://www.esi.dz>) Email address: (gm_abdelhamid, s_batata, f_azouaou) @esi.dz <https://arxiv.org/ftp/arxiv/papers/2201/2201.03423.pdf>

M. A and R. I, V. Palade, S. Amin and V. Chang “An Integrated Approach for Intrinsic Plagiarism Detection” in 2019 Future Generation Computer

(ELSEVIER) volume 96, Page 700-712.

<https://www.sciencedirect.com/science/article/abs/pii/S0167739X17326018>

A. S. Bin-Habtoor and M. A. Zaher “A Survey on Plagiarism Detection System” in 2012 International Journal of Computer Theory and Engineering Vol. 4, No. 2.

<https://www.researchgate.net/profile/Mahmoud-Zaher->

[2/publication/271302675_A_Survey_on_Plagiarism_Detection_Systems/links/62163279ba15e05e2ea381eb/A-Survey-on-Plagiarism-Detection-Systems.pdf](https://doi.org/10.1177/2158244015590162)

Salim Razi “Development of a Rubric to Assess Academic Writing Incorporating Plagiarism Detectors”(RS) SAGE

Open (2015),[10.1177/2158244015590162](https://doi.org/10.1177/2158244015590162)

<https://journals.sagepub.com/doi/pdf/10.1177/2158244015590162>

Debora Weber-Wulff “Plagiarism detectors are a crutch, and a problem” March 27, 2019

Nature (2019),[10.1038/d41586-019-00893-5](https://doi.org/10.1038/d41586-019-00893-5) <https://www.nature.com/articles/d41586-019-00893-5>

Sharma, S., Sharma, C. S., & Tyagi, V. (2015). Plagiarism detection tool “parikshak.” In Proceedings - 2015 International Conference on Communication, Information and Computing Technology, ICCICT 2015. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICCICT.2015.7045739>

Karnalim, O., Simon, & Chivers, W. (2019). Similarity Detection Techniques for Academic Source Code Plagiarism and Collusion: A Review. In TALE 2019 - 2019 IEEE International Conference on Engineering, Technology and Education. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/TALE48000.2019.9225953>

Steven Bradley “Creative assessment in programming: Diversity and divergence” in 2020 ACM International Conference Proceeding Series Article No.: 13Pages 1–4 <https://doi.org/10.1145/3372356.3372369>

- Gomes, K. P., & Matos, S. (2019). Contributions of Bioinformatics for Computing Education in the Detection of Programming Assignment Plagiarism (p. 1351). Sociedade Brasileira de Computacao - SB. <https://doi.org/10.5753/cbie.sbie.2019.1351>
- Nguyen-Son, H. Q., Thao, T. P., Hidano, S., & Kiyomoto, S. (2019). Detecting machine-translated text using back translation. In INLG 2019 - 12th International Conference on Natural Language Generation, Proceedings of the Conference (pp. 189–197). Association for Computational Linguistics (ACL). <https://doi.org/10.18653/v1/w19-8626>
- Atadjanov*, J., & Atadjanov, B. (2019). Cross-Language Plagiarism Detection Based on CLAD Method. International Journal of Innovative Technology and Exploring Engineering, 9(2), 4903–4909. <https://doi.org/10.35940/ijitee.b7404.129219>
- English, R. (2019). Designing computer security assessments to reduce plagiarism. In ACM International Conference Proceeding Series. Association for Computing Machinery. <https://doi.org/10.1145/3294016.3294020>
- Hourrane, O., & Benlahmer, E. H. (2019). Rich style embedding for intrinsic plagiarism detection. International Journal of Advanced Computer Science and Applications, 10(11), 646–651. <https://doi.org/10.14569/IJACSA.2019.0101185>
- Bayudan, J. H. C., Pati, F., Slater, R. E., & Palaoag, T. D. (2020). Safe art: A digital artwork plagiarism detector. In Theory and Practice of Computation (pp. 167–170). CRC Press. <https://doi.org/10.1201/9780429261350-18>
- Deng, P., Linsky, C., & Wright, M. (2020). Weaponizing Unicodes with Deep Learning - Identifying Homoglyphs with Weakly Labeled Data. In Proceedings - 2020 IEEE

International Conference on Intelligence and Security Informatics, ISI 2020. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ISI49825.2020.9280538>

Bhowmik, M., Nara, M., & Mohan, B. R. (2020). Similarity Calculation of Executable Using Intel Pin Instrumentation Framework. In Proceedings - 2020 IEEE 31st International Symposium on Software Reliability Engineering Workshops, ISSREW 2020 (pp. 169–170). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ISSREW51248.2020.00066>

Chavan, H., Taufik, M., Kadave, R., & Chandra, N. (2021). Plagiarism Detector Using Machine Learning. International Journal of Research in Engineering, Science and Management, 4(4), 152–154. Retrieved from <https://www.ijresm.com>

Horn, M. (2015). Postmodern Plagiarisms. Postmodern Plagiarisms. DE GRUYTER. <https://doi.org/10.1515/9783110379105>

Tripathi, R., Tiwari, P., & Nithyanandam, K. (2015). Avoiding plagiarism in research through free online plagiarism tools. In 2015 4th International Symposium on Emerging Trends and Technologies in Libraries and Information Services, ETT LIS 2015 - Proceedings (pp. 275–280). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ETTLIS.2015.7048211>

Related Studies

Jahangard, S. (2018). Exploring rotation and scale invariant features in image plagiarism detection using manifold-ranking algorithm. International Journal of Engineering and Technology (UAE), 7(4), 2663–2671. <https://doi.org/10.14419/ijet.v7i4.10575>

Vicent Naano Anney & Mary Atanas Mosha “Student's Plagiarisms in Higher Learning Institutions in the Era of Improved Internet Access: Case Study of Developing Countries” in 2015 Journal of Education and Practice ISSN 2222-1735 (Paper) ISSN 2222-288X (Online) Vol.6, No.13 <https://files.eric.ed.gov/fulltext/EJ1080502.pdf>

Rodriguez-Torrejón, D. A., & Ramos, J. M. M. (2014). CoReMo 2.3 Plagiarism Detector Text Alignment Module. Working Notes of the Conference and Labs of the Evaluation Forum (CLEF), 1180, 997–1003. Retrieved from <http://ceur-ws.org/Vol-1180/CLEF2014wn-Pan-RodriguezTorrejonEt2014.pdf>

Ratna, A. A. P., Purnamasari, P. D., Adhi, B. A., Ekadiyanto, F. A., Salman, M., Mardiyah, M., & Winata, D. J. (2017). Cross-language plagiarism detection system using latent semantic analysis and learning vector quantization. Algorithms, 10(2). <https://doi.org/10.3390/a10020069>

Paymode, N., Yadav, R., Vichare, S., & Bhoir, S. (2021). Online Assignment Plagiarism Detector. International Journal of Advanced Research in Science, Communication and Technology, 528–535. <https://doi.org/10.48175/ijarsct-1057>

Yalcin, K., Cicekli, I., & Ercan, G. (2022). An external plagiarism detection system based on part-of-speech (POS) tag n-grams and word embedding. Expert Systems with Applications, 197. <https://doi.org/10.1016/j.eswa.2022.116677>

Selemani, A., Chawinga, W. D., & Dube, G. (2018). Why do postgraduate students commit plagiarism? An empirical study. International Journal for Educational Integrity, 14(1). <https://doi.org/10.1007/s40979-018-0029-6>

- Levine, J., & Pazdernik, V. (2018). Evaluation of a four-prong anti-plagiarism program and the incidence of plagiarism: a five-year retrospective study. *Assessment and Evaluation in Higher Education*, 43(7), 1094–1105. <https://doi.org/10.1080/02602938.2018.1434127>
- Denney, V., Dixon, Z., Gupta, A., & Hulphers, E. (2021). Exploring the Perceived Spectrum of Plagiarism: A Case Study of Online Learning. *Journal of Academic Ethics*, 19(2), 187–210. <https://doi.org/10.1007/s10805-020-09364-3>
- Coughlin, P. E. (2015). Plagiarism in five universities in Mozambique: Magnitude, detection techniques, and control measures. *International Journal for Educational Integrity*, 11(1). <https://doi.org/10.1007/s40979-015-0003-5>
- Sibomana, E., Ndayambaje, I., & Uwambayinema, E. (2018). Plagiarism in higher education environment: causes and solutions. *Rwandan Journal of Education*, 4(2), 15–23. <https://www.mendeley.com/catalogue/b3916ed3-071c-3152-a4b3-4078e21eaf99/>
- Callahan, J. L. (2018). The retrospective (im)moralization of self-plagiarism: Power interests in the social construction of new norms for publishing. *Organization*, 25(3), 305–319. <https://doi.org/10.1177/1350508417734926>
- Bouarara, H. A., Hamou, R. M., & Rahmani, A. (2017). BHA2: Bio-inspired algorithm and automatic summarisation for detecting different types of plagiarism. *International Journal of Swarm Intelligence Research*, 8(1), 30–53. <https://doi.org/10.4018/IJSIR.2017010102>

Reyneke, Y., Shuttleworth, C. C., & Visagie, R. G. (2021). Pivot to online in a post-COVID-19 world: critically applying BSCS 5E to enhance plagiarism awareness of accounting students. *Accounting Education*, 30(1), 1–21. <https://doi.org/10.1080/09639284.2020.1867875>

Alvi, F., Stevenson, M., & Clough, P. (2021). Paraphrase type identification for plagiarism detection using contexts and word embeddings. *International Journal of Educational Technology in Higher Education*, 18(1). <https://doi.org/10.1186/s41239-021-00277-8>