PAMANTASAN NG LUNGSOD NG MUNTINLUPA

A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects

A Capstone Project Presented to the Faculty of the
College of Information Technology and Computer Studies
Pamantasan ng Lungsod ng Muntinlupa

In Partial Fulfilment of the Requirements for the Degree
Bachelor of Science in Information Technology

MALONZO, ALYSSA JHARIE A.
ECHEVARRIA, REMELYN B.
ULNAGAN, ROSE MARIE B.

JUNE 2025

PAMANTASAN NG LUNGSOD NG MUNTINLUPA

TABLE OF CONTENT

List of Tables

List of Figures

Chapter 1
INTRODUCTION

1.1 PROJECT CONTEXT

The proposed project, titled "A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects," will be implemented at Pamantasan ng Lungsod ng Muntinlupa (PLMun). The primary stakeholders and beneficiaries of this system are CITCS College of Information Technology and Computer Studies and the University Library. The development and deployment of the system will leverage the university's existing digital infrastructure, including servers and network facilities, to ensure reliable access, security, and scalability. The system is designed to be accessible to students, faculty, administrative staff, and library personnel within the campus, providing a centralized digital repository for academic outputs. This platform aims to promote efficiency, facilitate easy access to research outputs, and enhance the management, preservation, and accessibility of thesis and capstone projects.

The current system at PLMun faces several challenges. Physical storage of thesis and capstone projects makes access cumbersome, especially for users who cannot physically visit the library or storage areas. Digital files are often disorganized or stored in multiple locations, making retrieval time-consuming and inefficient. Existing search methods rely on basic keywords, which limit the precision and speed in finding relevant documents, especially as the volume of stored works increases. Furthermore, there is no integrated system to verify the originality of submitted works, raising concerns about academic integrity and potential plagiarism. Manual management of these documents is prone to errors, inconsistencies, and difficulties in updating or maintaining the repository. These issues highlight the urgent need for an organized, intelligent, and secure digital archive that can streamline access, improve search capabilities, and uphold academic standards.

Addressing these challenges is essential for advancing the academic environment at Pamantasan ng Lungsod ng Muntinlupa. Developing an intelligent, web-based archive will enhance accessibility to thesis and capstone projects across the university community, regardless of location, thereby fostering a culture of openness and knowledge sharing. Integrating AI-powered search capabilities will enable users to retrieve relevant documents efficiently and accurately, saving valuable time. The inclusion of an automated plagiarism detection tool will help uphold academic integrity by providing immediate feedback on originality, discouraging dishonest practices. Automating the management process will reduce manual errors, ensure data consistency, and facilitate easier updates and maintenance of the repository. Overall, this system aligns with the university's goal of embracing digital transformation, improving research quality, and promoting academic excellence, while serving as a vital platform for collaboration, innovation, and integrity among students and faculty.

Beneficiaries of this innovative system include students, who will gain easier access to a comprehensive database of previous research works for inspiration and citation, as well as prompt plagiarism checks for their submissions. Faculty members will benefit from streamlined processes for thesis review and approval, with tools to verify originality efficiently. Administrators will find it easier to organize, manage, and monitor the repository, ensuring data security and generating reports on academic integrity. Future researchers will have access to a rich and reliable source of scholarly works, encouraging further research and academic growth. Overall, this project aims to transform how academic outputs are stored, accessed, and validated at Pamantasan ng Lungsod ng Muntinlupa, fostering a more efficient, transparent, and integrity-driven scholarly environment.

1.2 PURPOSE AND DESCRIPTION

This study aims to create a Smart Web-Based Archive with AI-powered Search and Plagiarism Detection. It is specially designed for the university's thesis and capstone

projects at Pamantasan ng Lungsod ng Muntinlupa. The system will serve as a central, easy-to-use, and secure place for storing, finding, and checking academic work. This will help improve access to research, ensure the originality of projects, and support the overall quality of learning and research at the university.

Below are following beneficiaries who will benefited the system:

Company Staff/Personnel, They will be able to quickly find important research and project documents. This makes collaboration with the university easier and helps in making better decisions based on relevant research.

Researchers will enjoy an advanced AI search feature that helps them find related thesis and capstone projects easily. The built-in plagiarism check also ensures that the research is original and trustworthy, supporting credible work.

Future Researchers, Future students and researchers will have a reliable digital library to support their studies and research. The AI tools and plagiarism checks will help them review literature thoroughly and produce original projects, encouraging innovation and integrity.

1.3 OBJECTIVE OF THE STUDY

General Objective

The main objective of the study is to design and develop a Web-based Archiving System that features AI-enhanced search capabilities and an integrated plagiarism detection engine for managing and accessing completed theses and capstone projects within a university.

Specific Objectives

1. To design the system using the following features:
    a. Capable of Uploading and storing academic documents such as theses and capstone projects in a centralized web-based repository.

b. Capable of Performing intelligent search using AI, enabling users to search documents by keywords, authors, titles, topics, or abstracts.

c. Capable of Detecting potential plagiarism by comparing new submissions against existing documents stored in the system.

d. Capable of Detecting potential plagiarism by comparing new submissions against existing documents stored in the system.

e. Capable of Supporting a user-friendly interface to facilitate smooth navigation and interaction for all types of users.

2. To develop the system using PHP Programming, HTML, CSS, JavaScript and MySQL for the designing of the databases;
3. To test and improved the system using Alpha and Beta Testing method.
4. To evaluate the performance of the system using ISO/ IEC 25010: 2011 Software Characteristics.
5. To implement the system, a secure, scalable, and user-friendly web-based platform will be developed, integrating advanced AI-powered search algorithms and comprehensive plagiarism detection tools. The implementation process will involve designing a modular architecture utilizing modern web development frameworks, and conducting rigorous testing to ensure system reliability, security, and usability.

1.4 SCOPE AND LIMITATION

The proposed Web-Based Archive System for Pamantasan ng Lungsod ng Muntinlupa aims to create a user-friendly platform for the preservation, organization, and accessibility of students' theses and capstone projects within the university. It consists of two main modules: an Administrator module, which allows authorized personnel such as librarians and administrators to manage user accounts, review and

approve or reject uploads, and edit project details securely; and a User (Student/Researcher) module, enabling students to upload, view, download, and update their project information following a standardized journal-like format to protect intellectual property and provide comprehensive documentation. The system will be evaluated based on functionality, usability, reliability, and security to ensure it meets user requirements and maintains data integrity and privacy.

However, this study is limited to the College of Information Technology and Computer Studies, specifically targeting BSIT, BSCS, and MIT students. The migration of existing physical projects to the digital platform may require substantial time and resources. Additionally, communication with previous researchers—particularly those from 2019 and earlier—may be limited, potentially affecting the completeness of project data. Regular maintenance and monitoring are necessary to ensure smooth operation and data security. Overall, while the system aims to enhance project management and accessibility within the university, certain constraints related to infrastructure, resources, and external dependencies will influence its development and deployment.

Chapter 2

REVIEW OF RELATED LITERATURES & STUDIES

This chapter presents the review of related literature and studies. These pieces of literature and studies will be reviewed as they are deemed relevant to the present study. Researchers gathered information was congregated through articles, journals and books that are related with the study which serve as the foundation for collecting more knowledge and understanding the essence of the study for additional reference

2.1 TECHNICAL BACKGROUND

The development of "A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects" will utilize a combination of modern web technologies and tools to ensure a robust, efficient, and user-friendly system. Visual Studio Code (VSCode) will serve as the primary Integrated Development Environment (IDE), chosen for its versatility and support for multiple programming languages essential to the project. The front-end interface of the system will be built using HTML, CSS, and JavaScript to create an intuitive and responsive webpage that allows users to upload, view, and search thesis and capstone projects. PHP will be employed for server-side scripting to handle data processing, user authentication, and communication between the front-end and back-end databases.

The main file format used for files uploaded to the archive will be PDF, as it preserves the formatting and content integrity of academic documents. Metadata associated with each project such as Title, Author, Year, and Abstract will be stored in a MySQL database, enabling efficient management and retrieval of project information. To store the actual PDF files securely and reliably, Google Drive cloud storage will be integrated into the system, providing scalable and accessible file storage.

2.2 RELATED LITERATURE

[1] The Iloilo Science and Technology University Miagao Campus has created a system for storing and managing student research papers, which benefits both the research and development office and the university library. However, this system isn't accessible online and doesn't have a user-friendly interface. The aim of this study is to develop an online platform for archiving and profiling student theses, which will display research titles and abstracts categorized by year, course, author, and adviser, and allow users to save and retrieve data from a database. The development process used a prototyping approach within the software development lifecycle, following strict guidelines based on ISO 25010 quality standards. To evaluate the system, five research coordinators from different courses and ten IT experts were randomly selected. Results showed that the system performed very well across various quality attributes such as functionality, performance, compatibility, usability, reliability, security, maintainability, and portability, with an average score of 4.81 out of 5. In conclusion, the project successfully developed an online thesis archiving and profiling system that meets ISO 25010 standards, enhancing accessibility and ease of use for both the research office and the university library.

[2] This project focuses on developing and implementing the "Capstone Archiving Management System with Citation and Reference Generator" to address the challenges faced by School Library and Capstone Coordinators. Currently, CHMSU-FT relies on outdated methods for storing capstone projects, involving the tedious process of printing physical copies. This approach consumes a lot of paper and takes up valuable space in the school. The new system aims to improve the way capstone projects are archived and retrieved by providing a digital platform that allows quick and easy access to materials for students and faculty. It also helps reduce administrative tasks for coordinators and librarians, making the process more efficient and less labor-intensive.

Additionally, the system includes a citation and reference generator to simplify referencing tasks. This digitization promotes sustainability and modernizes record-keeping, offering a useful tool for future researchers to build upon. Based on feedback and data collected, the researchers concluded that the system is functional and beneficial. Users agree that it is useful for easily searching and citing capstone projects, and they find it simple to use for basic searches and referencing.

[3] This project introduces a web-based system designed to store and manage the studies, student profiles, and projects from Colon National High School's STEM program. The system was developed using technologies such as HTML, CSS, JavaScript, PhpMyAdmin, XAMPP, Visual Studio, and MySQL. To evaluate its performance, testing was carried out on devices including a Galaxy Tab A8, Oppo A9 2020, and an Acer Aspire E5-457G. The tests focused on assessing the system's speed and ease of access through Google search, as well as how user-friendly different parts of the website are. These evaluations involved STEM students, teachers, and IT experts. The performance tests were conducted using the best available Wi-Fi connection to ensure optimal conditions. However, the study has limitations, such as only testing with one Wi-Fi provider, a small number of devices, and testing in a single location.

[4] carried out a qualitative case study to explore how digital archiving and document management are organized at Taibah University. The focus was on meeting the needs of researchers, professors, administrators, and students, as well as ensuring that data is stored securely and accessible when needed. Using a case study approach allowed for a detailed analysis of the university's data management and archiving strategies. The findings showed that while Taibah University has a specific approach to managing electronic documents, it does not rely on a single, dedicated data management system. Instead, they archive only certain types of documents, and their data security measures align with university policies. Overall, the system they use is considered effective because it addresses the university's needs for secure data storage and

archiving. The university has established policies for collecting and preserving research and activity-related data of staff and students. Although their data management approach isn't a traditional digital system, it still successfully meets the institution's goals, as supported by the TTF model.

[5] was created as a key software tool to improve how research papers are stored and organized at the University of Industrial Technology (CIT). The system was carefully developed to be easy to use, with effective search features, the ability to track thesis and dissertation projects, and simplified processes for borrowing and returning books. The development followed a descriptive and developmental research approach, using a standardized method for collecting data. The software was built using the Rapid Application Development (RAD) model, which focuses on quick prototyping and iterative improvements to create a reliable system. The main users included CIT department administrators, faculty, students from CHMSU, and external researchers. To assess how well the system worked and how user-friendly it was, the Post-Study System Usability Questionnaire (PSSUQ) was used, resulting in a high average score of 2.28, indicating strong usability and positive reception among users. The results highlighted the system's high acceptance and usefulness, making it a valuable tool for the university community. Due to its improved accessibility, efficiency, and ability to quickly retrieve information, it was recommended that the system be adopted throughout the institution to enhance archiving and research processes for all users.

[6] uses the System Development Life Cycle (SDLC) approach to create a system that ensures data is processed in an integrated way and can be accessed by the community at any time. At Universitas Lancang Kuning (UNILAK), the Badan Penjaminan Mutu (BPM) is responsible for managing both internal and external quality assurance systems. Currently, they face challenges because they lack an IT-based system to handle the data related to internal quality assurance (SPMI) and accreditation, leading to delays and inefficiencies. To address this, the study recommends developing an

electronic document management system that can process accreditation data faster, more accurately, and in a more integrated fashion. The goal is to improve how information is presented and make the accreditation data more accessible and reliable for the academic community through a computerized system.

[7] adopted a system development research approach that involved two main stages: first, analyzing the needs, and second, designing and developing the record archiving system. Proper records management and storage are crucial for running the university efficiently and effectively, as they help document planning and service implementation while enabling proper tracking of activities. In this study, the researcher created a digital record archiving system for Eastern Visayas State University Burauen Campus. This system aims to eliminate paper-based records, making it easier and faster to retrieve and store data, while also ensuring a secure backup of the database. Currently, EVSU-BC relies heavily on physical records stored in steel cabinets, managed by the Records Team, which manually determines which records should be kept. The existing system uses multiple storage racks filled with various files, taking up a lot of space. Because the campus handles many records in different formats—both paper and electronic—the current archiving process is outdated, slow, and inefficient. Despite efforts in other universities to modernize record management through database systems, many still rely on traditional paper-based methods, leading to complaints from staff and issues such as insufficient qualified personnel and facilities. Therefore, the study aimed to develop a digital record archiving system that would provide a paperless management solution, streamline data retrieval and entry, and ensure a reliable backup of records.

[8] The Institute of Information and Computer Studies (IICS) faced a problem with faculty members having to submit documents at the end of each academic year, which are required by the Director of IICS before approving faculty clearances. To address this issue and speed up the process, a solution called the “e-Document

Archiving System with SMS Support" was developed. The study used descriptive research to evaluate whether the system met its goals, and developmental research to build the software itself. The system's architecture was based on a multi-tier (n-tier) design. Since the system would be installed on the local area network (LAN) within the IICS building, faculty members could upload and store their documents electronically. The Director could then easily download and print these documents in real time. Additionally, an SMS support feature was included to send notifications to faculty members about which documents they need to submit and the deadlines for submission. The results showed that uploading and downloading documents could be done instantly, making the process accessible in real time. The system was found to be user-friendly, effective, and efficient, meeting the needs of the users well.

[9] This paper discusses the development of a Smart Final Year Project Archive System (SFYPAS) built with the Laravel framework to help manage and access final year project (FYP) theses at Universiti Teknologi MARA (UiTM). The goal of the system is to create an online platform where students, supervisors, and lecturers can easily access previous FYP theses, making it easier than physically searching through thesis rooms. The system was developed following the Software Development Life Cycle (SDLC) method. To evaluate its effectiveness and user acceptance, the researchers used the Technology Acceptance Model (TAM). The results showed that users were generally satisfied, with an average score of 4.27 for ease of use and 4.47 for usefulness. Overall, the study found that SFYPAS successfully met its goals and serves as a useful platform for accessing past FYP theses.

In the context of UiTM, completing a final year project is a mandatory part of graduating from any bachelor's program. However, finding a suitable project topic can be difficult, and having access to previous theses can inspire ideas and help ensure the proposed project isn't a duplication of previous work. Unfortunately, these theses are usually stored in a physical room, which makes it inconvenient and time-consuming to search for references. The lack of an electronic management system for these documents

creates issues with organization and retrieval. To address this, the project aimed to develop SFYPAS using Laravel, applying SDLC in the process, and to evaluate its usability through TAM and expert heuristic evaluation. The system was tested with thirty users who provided feedback via questionnaires. The results confirmed high satisfaction levels, with perceived ease of use and usefulness scores indicating the system's effectiveness. In conclusion, the research achieved its goal of providing a web-based archive system that simplifies accessing past FYP theses. Future enhancements could include adding a search engine with algorithms to improve search results.

[10] This article aims to conduct a detailed analysis of how the city council manages its document flow and to create models that illustrate the internal processes involved in electronic document circulation once such systems are adopted within the organization. The research methodology includes reviewing existing scientific studies on electronic archives and methods for storing electronic documents, analyzing the current status of the city council's adoption of information technology in its workflows, selecting appropriate software tools, and examining the financial capacity of the institution along with the potential for using electronic solutions in document management. The innovative aspect of the study is its contribution to a broader understanding of the importance of modernizing document workflows in Ukrainian government bodies, especially city councils. This involves exploring the potential benefits of such innovations, providing a detailed look at how electronic archives operate internally, and choosing suitable technological options to achieve these goals.
moving to electronic archives can accelerate document processing and reduce expenses related to maintaining physical storage. As Ukraine advances in digital transformation, online document management solutions are becoming increasingly popular. However, some areas, particularly smaller towns and rural institutions, still face challenges. Implementing professional electronic archive systems can be costly, but if there are limited funds, free alternatives like Google Sheets can be used. Electronic archives will

enable institutions to work more efficiently and securely, especially during times of conflict, since employees can access documents remotely from any location with internet connectivity.

2.3 RELATED STUDIES

This study investigates whether cloud computing systems are adequate for managing the full life cycle of electronic records and archives, including maintaining their original order and provenance. The results show that there is no clear agreement on what constitutes records and archives architecture that effectively supports records management. The popularity of cloud computing services significantly influences how electronic records and archives are handled. Because employees often use multiple cloud platforms, organizational records are stored across various locations, which can violate the established principles, rules, regulations, and legal requirements for managing public records and archives. Therefore, safeguarding national and organizational secrets, as well as ensuring the long-term preservation of organizational memory in the cloud, has become a critical concern. As the use of different digital architectures for government services continues to grow, it is essential to study how cloud computing impacts the core principles of records and archives management. [11]

This study introduced a web-based platform designed to gather and store the activities of employees at an Iraqi university, utilizing both frontend and backend technologies, specifically the Node.js framework. The system includes different user roles, from administrators to employees, with roles such as the dean and department head also involved. In today's digital age, electronic archiving of employee activities is becoming increasingly important for organizations. It helps ensure accurate documentation of work, supports performance evaluations, highlights areas for improvement, and ensures compliance with policies and regulations. Employee activities are vital for enhancing the academic and societal contributions of institutions and universities by fostering participation, innovation, and development in various

fields. This system also plays a key role in assessing employee performance and their eligibility for promotions. The main advantage of this system is that it replaces manual searching and paper-based archiving with a quick, reliable electronic process that can be accessed and used from anywhere with an internet connection. [12]

This study focused on designing and developing an Electronic Document Archiving System for the University of Rizal System. The system helps users organize and access digital files easily. The system was tested by both teaching and non-teaching staff at URS Binangonan campus. To evaluate its performance, a survey questionnaire based on ISO 25010 software quality standards—covering aspects like functionality, performance, usability, reliability, security, maintainability, and portability—was distributed through Google Forms. The development process used the Rapid Application Development methodology within the Software Development Life Cycle. The survey results showed that the system's ratings in these areas ranged from acceptable to good, with average scores between 3.31 and 3.49 on a scale where higher scores indicate better performance. The study concluded that the system's performance is acceptable but recommended improvements, especially since IT experts, faculty, staff, and administrators identified weaknesses in areas such as functionality, efficiency, usability, reliability, maintainability, security, and portability based on ISO standards. [13]

The purpose of the Admission and Registrar Office (ARO) Mobile Document Scanner with Archiving System is to assist the registrar's office in managing student requests, admissions, and the handling of documents submitted during enrollment. This system aims to revolutionize how the office manages records by improving efficiency, security, and transparency. It will convert traditional paper documents—such as transcripts, enrollment forms, and other important records—into digital formats that can be accessed easily. To assess how well the system works, specific performance criteria were set. The system uses a star topology for communication interfaces. Feedback from 30 respondents evaluated features like accuracy, ease of use, reliability, operability,

learning support, interactivity, compatibility, and cross-platform functionality, all of which received an “excellent” rating. Overall, the system was designed to make the entire process more convenient and efficient for the university’s administrators and staff. [14]

This paper explores how integrating Geographic Information Systems (GIS) with open-source technologies can improve teaching and learning in the Geography and GIS departments at Cairo University. Recognizing the growing importance of technology in education, especially after the COVID-19 pandemic, the study introduces an online map archiving system. This system allows users to quickly search, view, and access digital maps from various sources through different query options, making the teaching process more efficient and encouraging greater student engagement with maps. The system’s effectiveness was assessed using checklists and questionnaires. Results show that the system significantly enhances learning by providing essential tools and creating a more interactive and secure environment. According to the authors, this is the first time such a GIS-based open-source map archiving system has been developed in Egyptian universities. Overall, the study highlights the potential of GIS and open-source tech to transform traditional teaching methods, fostering a more dynamic and technology-driven learning experience. [15]

This paper discusses the drawbacks of traditional methods used to supervise and manage final-year student projects. These conventional approaches are often slow, stressful, and expensive, involving manual steps like selecting topics, submitting proposals, defending projects, and archiving documents. To address these issues, a user-friendly and efficient web-based system was created for assigning, managing, and storing student projects. The system allows for real-time communication between supervisors and students, simplifies approval and submission procedures, and provides a central location for all project-related files. It was built using PHP for server-side functions, ASP.NET for the user interface, and MySQL for database management, all

integrated within XAMPP for local server operation on Windows. Testing shows that the system performs well and has the potential to replace manual methods in higher education. Overall, it helps reduce the effort, time, and challenges involved in supervising and managing final-year student projects. [16]

This study aimed to review the development of library and information science education in several Jordanian universities from its beginning in 1921 up to 2022. The focus was on analyzing how these programs have evolved and the challenges they face, particularly in the Bachelor's and Master's degrees offered by the Department of Library and Information Technology (DLIT) at Al-Hussein Bin Talal University. The researchers used descriptive and content analysis methods to examine the early days of establishing the first professional library and information science program in the late 1970s, and to evaluate the development of the recent master's program in Information Management and Digital Archiving at Al-Hussein Bin Talal University. The findings revealed that the four universities in Jordan lack coordination and collaboration in updating their curricula and teaching methods. Moreover, there is a deficiency of recent research on the importance of these issues within the higher education system. The study emphasizes the need for increased cooperation among faculty members across these universities to assess current conditions and develop more effective strategies to improve education in libraries, information science, documentation, and digital archiving. The fields are continuously evolving in Jordan, and the Department at Al-Hussein Bin Talal University is making notable progress to keep pace with advancements in information technology. The main conclusion is that academic programs should aim to bridge the gap between the skills needed by library professionals and the latest developments in technology, telecommunications, and digital transformation. [17]

This research focuses on the idea of archiving, which involves storing and providing access to past works. As technology advances, many universities have adopted electronic archive systems to store research projects, making it easier to access

these works and enhancing research productivity. However, Bowen University in Nigeria currently relies on a manual system for archiving projects, which is outdated compared to modern digital archiving methods. To address this, the study aimed to create an electronic project archive system for Bowen University that securely stores past projects and makes them accessible to support undergraduate research and contribute to knowledge sharing. The system, called Bowen University Project Archive System (BUPAS), was developed using HTML for the user interface, PHP to handle system operations and database interactions, and WAMP server for local hosting, enabling smooth communication between components. Usability testing was performed to gather feedback from users. The development followed an agile methodology, resulting in a fully functional system that offers a faster, simpler way for students to submit and access projects. Users appreciated the system's ease of use and straightforward navigation. An important feature of any archive system is security to ensure the authenticity and integrity of the stored materials, and BUPAS successfully incorporates this security aspect. [18]

This study introduces an integrated platform that combines an Institutional Repository (IR) with a Learning Management System (LMS) to improve the educational experience at Higher Education Institutions (HEIs). Traditionally, these functions are handled by separate systems—one for making institutional outputs available (IR) and another for supporting ICT-based learning (LMS). The proposed solution merges these functions into a single, unified platform, providing a smooth experience for both faculty and students. The platform includes features like access to a Remote Lab and simulations based on Scilab. Users log in with a single ID, making it easy to switch between different parts of the system. Course materials and other learning resources are stored in the IR but are referenced within the LMS, which helps reuse content, reduces duplication, and simplifies updates. Student activities on the LMS are linked to content access, and the platform offers a large collection of open access resources. Now in its fourth version and operational for over 20 years, the system contains more than 20,000

digital resources, including nearly 250 open access learning materials. The main goal is to unify LMS and IR functions to improve education and research. The paper provides examples from two electrical engineering courses—Signals and Systems, and Electric & Electronic Circuits—to illustrate how the system works. [19]

Institutional repositories (IRs) have gained significant attention from researchers worldwide across various disciplines. They are valuable for increasing the public impact, reputation, and visibility of both researchers and universities. Despite the rapid growth and importance of IRs, there has been little systematic review or integration of existing research findings, or an assessment of the current state of knowledge about IRs. The main purpose of this paper is to offer a comprehensive understanding and detailed review of the current research related to IRs. The study used a systematic literature review (SLR) approach, following a structured protocol to organize the research effectively. Data was collected from primary studies published between 2007 and 2018 across six major databases: ScienceDirect, IEEE Explorer, Springer, ACM, Taylor and Francis, and Emerald Insight. After applying specific inclusion and exclusion criteria, 115 studies were selected for analysis. The review revealed that a major challenge in developing open access IRs is the lack of awareness among scholars and institutions about open access repositories, along with inadequate ICT infrastructure. On the positive side, IRs offer benefits such as increased visibility for academic institutions, higher local and global rankings, enhanced prestige and public value, and support for better teaching, learning, and research activities. The review also found that most studies focus on the deployment, implementation, adoption, benefits, and challenges of IRs. The findings from this research can guide future studies by providing a clear overview of IRs and offering recommendations for successful implementation in higher education institutions. [20]

2.4 DEFINITION OF TERMS

OPERATIONAL TERMS

1. Accreditation Data

Data related to the process of officially recognizing the standards and quality of an institution or program.

2. Administrator

A user role responsible for managing and overseeing the system's operations and user access.

3. Data Security

Measures implemented to protect data from unauthorized access or breaches.

4. Data Management

The organization, storage, and maintenance of data within a system or institution.

5. Efficiency

The ability of the system or process to perform its functions with minimal waste of time and resources.

6. Evaluation

The process of assessing the system's performance, usability, and effectiveness.

7. File/Document Storage and Retrieval

The process of saving digital or physical files for future access and management.

8. Information Accessibility

The ease with which users can find and use relevant data or documents.

9. Institutional Repository (IR)

A digital platform for storing, managing, and providing access to an institution's scholarly outputs and research materials.

10. Management of Student Research Papers/Projects

Organizing, storing, and overseeing student academic work, including theses, capstone projects, or research papers.

11. Monitoring and Tracking

Observing and recording the progress or status of documents, projects, or activities.

12. Organization and Management

The arrangement and administration of resources, data, or activities within a system or institution.

13. Security Policies

Rules and procedures to protect data and ensure system security.

14. User Roles and Permissions

Defined levels of access granted to different users based on their responsibilities.

TECHNICAL TERMS

1. HTML (Hypertext Markup Language)**

The standard markup language for creating web pages and web applications.

2. ISO 25010

An international standard for software product quality, covering attributes such as functionality, usability, reliability, security, maintainability, and portability.

3. Laravel Framework

An open-source PHP web framework used for developing web applications.

4. MySQL

An open-source relational database management system.

5. Node.js

An open-source JavaScript runtime environment used for developing server-side applications.

6. PHP (Hypertext Preprocessor)

A server-side scripting language used to develop dynamic websites and applications.

7. Rapid Application Development (RAD)

An iterative software development methodology emphasizing quick prototyping and user feedback.

8. System Development Life Cycle (SDLC)

A structured approach to developing information systems through phases like planning,

analysis, design, implementation, and maintenance.

9. Software Quality Standards (ISO 25010)

Criteria for evaluating the quality of software products.

10. Technology Acceptance Model (TAM)

A framework for assessing user acceptance and perceived usefulness of a technology.

11. WAMP Server (Windows, Apache, MySQL, PHP)

A Windows-based platform for developing and testing web applications.

12. Wired/Wireless Networks

Network types used for data transmission, including LAN and Wi-Fi.

13. Web-based System

An application accessible via internet browsers.

14. WAMP Server

A platform for local web development combining Windows, Apache, MySQL, and PHP.

REFERENCES

[1] Montalban, L., et al., (2023) "ISAT U Miagao Campus Students Thesis Archiving and Profiling System" International Journal of Computer Science and Information Technology Research, 11 (2), 50-54

[2] Doromal, W., & Soberano, K. (2023) "Capstone Archiving Management System with Citation and Reference Generator" International Journal of Research in Engineering and Science (IJRES), 11(9), 128-133

[3] Cagang, T., et al. (2023) "Stem Web-Based Sysmtem For Student's Profiling And Archiving Of Thesis And Capstone Projects" International Journal of Novel Research and Development (IJNRD), 8(6), 493-532

[4] Alokluk, J., et al. (2019) "Archiving and Document Management at Taibah University: A Case Study" Canadian Center of Science and Education, 12 (4), 11-19. DOI:10.5539/cis.v12n4p11

[5] Mondero, N,. et al. (2024) "Preserving Scholarly Endeavors: A System for Archiving Thesis and Dissertations" GAS Journal of Multidisciplinary Studies (GASJMS), 02 (04), 44-52. DOI:10.5281/zenodo.11063772

[6] Febriadi, B,. & Rahmat, A,. (2019) "Effectiveness of E-Archives As A Media Of Information Technology Based Data Management In Universitas Lancang Kuning" International Journal of Information System & Technology, 04 (01), 502-508.

[7] Villarosa, E,. (2021) "DEVELOPING A RECORD ARCHIVING SYSTEM IN EASTERN VISAYAS STATE UNIVERSITY – BURAUEN CAMPUS " International Journal of Multidisciplinary Academic Research, 9 (3), 64-73.

[8] Alpasan, B,. et al,. (2021) "E-Document Archiving System with SMS Support" INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND ANALYSIS, Volume 04 Issue 12, 1935-1943. DOI : https://doi.org/10.47191/ijmra/v4-i12-24

[9] MatTaib, A., Mohd Shukri, N. A. ., Ahmad Zukri, N. H., & Ghazali, N. (2020). "Smart Final Year Project Archive System Using Laravel Framework with Email Notification (SFYPAS)." Journal of Computing Research and Innovation, 5(2), 41–53. DOI: https://doi.org/10.24191/jcrinn.v5i2.145

[10] Ruslan, K,. et al,. (2023) "Electronic archive as a means of quick access to management information" Library Science. Document Science. Informatology, Vol. 19, No. 4, 14-21. DOI: 10.32461/2409-9805.4.2023.293967

[11] Aisyah Katrina Julaihi, Siti Noorsiah Jamaludin, Nor Kamariah Chik, Rusnah Johare, "Digital Record-Keeping Practices: Electronic Records and Archives in the Cloud," International Journal of Engineering Trends and Technology,vol. 72, no. 10, pp. 267-281, 2024. Crossref, DOI: 10.14445/22315381/IJETT-V72I10P126

[12] Taha Adel Taha,. Ayad Hussain Abdulqader,.(Jul 2023)"Web-Based System for Archiving Activities of University College Staff" International Research Journal of Innovations in Engineering and Technology, Vol. 7, Iss. 7, 113-117. DOI:10.47001/IRJIET/2023.70701

[13] Daluyon, M,. (2024)"Design and Development of Electronic Document Archiving System", Journal of Innovative Technology Convergence,6 (4), DOI: 10.69478/JITC2024v6n4a03.

[14] Mallares, D., Alegata, W.C. and Bual, J. (2024)." Admission and Registrar Office (ARO) Mobile Document Scanner with Archiving System." Technium: Romanian Journal of Applied Sciences and Technology. 22, (Apr. 2024), 33–54. DOI:https://doi.org/10.47577/technium.v22i.10957.

[15] Bedair, S,. et al,. (2022) "Enhancing Hybrid Learning using Open Source GIS-Based Maps Archiving System" The Egyptian Journal of Remote Sensing and Space Sciences, Volume 25, Issue 3, Pages 779-793. DOI: https://doi.org/10.1016/j.ejrs.2022.07.003

[16] Adamu, A,. (2020) "FINAL YEAR STUDENT PROJECT ALLOCATION ARCHIVING AND MANAGEMENT SYSTEM" Tertiary Education Trust Fund, Vol. 5 No. 3 pp. 837 – 840.

[17] AL TAWALBEH, S. A., SULEIMAN, R. J. ., & AHMED, M. H. (2024). “Developments of Teaching Library, Information and Digital Archiving Sciences in Jordan: Al-Hussein Bin Talal University as a Case Study”. University Library at a New Stage of Social Communications Development. Conference Proceedings, (9), 299–315. DOI: https://doi.org/10.15802/unilib/2024_317501

[18] Okedigba, Temilola O., Ibukun Okedigba, and Hallelujah O. Aworinde.(2020) “Research Archiving As a Tool for Research Continuity: A Look at Bowen University, Nigeria”. Asian Journal of Research in Computer Science 5 (3):52-62. DOI: https://doi.org/10.9734/ajrcos/2020/v5i330139.

[19] G. Penello Temporão, et al,. (2019) “THE INTEGRATION OF AN INSTITUTIONAL REPOSITORY AND A LEARNING MANAGEMENT SYSTEM: A CASE STUDY” International Technology, Education and Development Conference, 3314-3320. DOI: 10.21125/inted.2019.0862

[20] S. Asadi, R. Abdullah, Y. Yah and S. Nazir, "Understanding Institutional Repository in Higher Learning Institutions: A Systematic Literature Review and Directions for Future Research," in IEEE Access, vol. 7, pp. 35242-35263, 2019, DOI: 10.1109/ACCESS.2019.2897729

CHAPTER 3

METHODOLOGY

In this chapter, the methodologies used throughout the study are described together with the diagram presentation. Here are the following sections that will be discussed in this chapter: Requirements Analysis, Requirement Documentation, Design of Software, System Product and/or Process, Development and Testing, and Implementation Plan for the project titled "A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects."

3.1 REQUIREMENTS ANALYSIS

In this section,In this phase, show the details of the current system function: who (the people who are involved), what (the business activity), where (the environment in which the work takes), when (the timing), and how (how the current procedures are performed) of the business or organization understudy

Who (the people who are involved)

Researchers - Responsible for conducting research activities and preparing their thesis or capstone projects for submission into the archive.

Students - The primary users who upload their completed works, access stored documents for reference, and utilize the archive for research and review purposes.

Librarians - Oversee the organization, cataloging, and management of the digital repository, ensuring efficient retrieval and proper maintenance of the archived documents.

Subject Professors - Serve as academic reviewers and approvers of thesis and capstone submissions, ensuring compliance with academic standards and quality assurance.

Technical Advisers - Provide guidance and support to students and researchers throughout the research process and submission procedures.

IT Experts - Responsible for the technical maintenance of the web-based platform, implementing AI-powered search and plagiarism detection functionalities, and ensuring system security, reliability, and performance.

Great! Here's the continuation for the "What," "Where," "When," and "How" components, aligned with the title and context of your project:

What (The Business Activity)

The primary activity involves the collection, submission, organization, and retrieval of thesis and capstone projects within the university's digital archive. This includes students submitting their completed works, librarians cataloging and maintaining the digital repository, and users accessing stored documents for research, review, or verification purposes. The system also incorporates AI-powered search functionalities to facilitate quick and accurate retrieval of documents, as well as plagiarism detection to ensure academic integrity.

Where (The Environment in which the Work Takes Place)

The activities are conducted within a university environment, utilizing a web-based platform accessible via the internet. The digital archive is hosted on university servers or cloud infrastructure, allowing authorized users, students, faculty, librarians, and administrators to access the system remotely from various locations within the campus or off-campus with internet access.

When (The Timing of Business Activities)

The processes occur throughout the academic calendar, especially during thesis and capstone submission periods. Students upload their projects upon completion, and librarians or faculty review submissions within specified deadlines.

How (The Procedures and Processes)

Current Technical Process

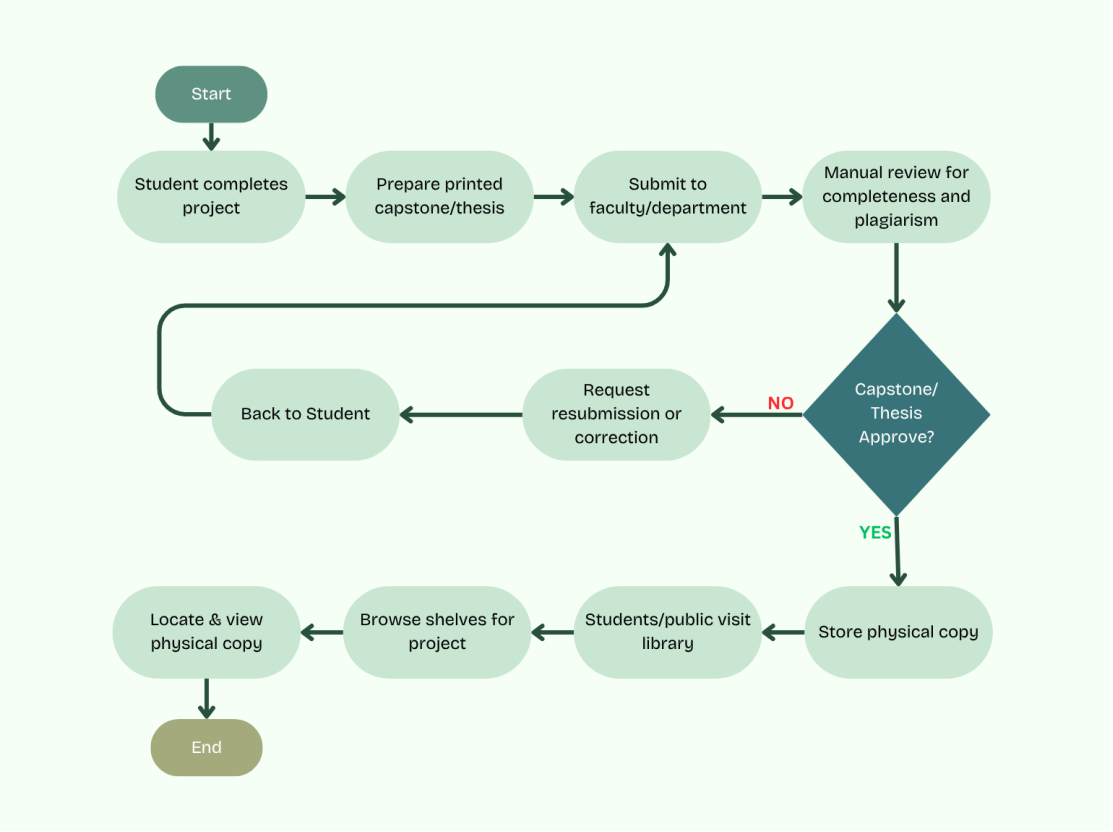

Figure 1. Current Technical Process for Archiving System

The figure above depicts the current data archiving workflow for university projects. It illustrates the process from project completion by students to storage and retrieval. The flowchart outlines key steps: preparing materials, submission, review, cataloging, and physical shelving. It also shows how users access projects through searching the catalog and locating physical copies.

3.2 REQUIREMENT DOCUMENTATION

User Requirements

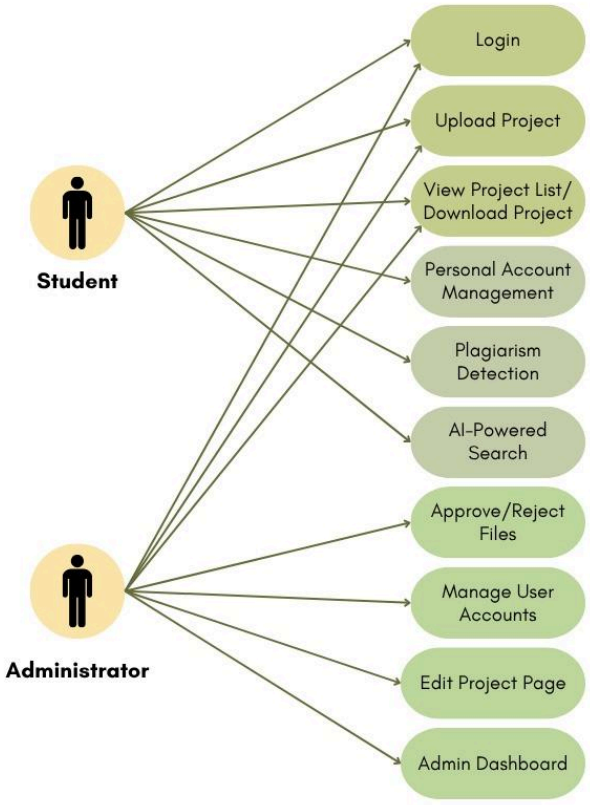

Figure 2. Use Case Diagram of A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects.

The figure above presents the Use Case Diagram of the Web-Based Capstone and Thesis Archiving System for PLMUN-CITCS. It illustrates how different users interact with the system to perform various tasks. There are two primary user roles: the administrator and the student researcher. The administrator has a broader set of capabilities, including managing and maintaining the system, as well as overseeing user activities. Meanwhile, the student researcher can access and upload their capstone or thesis documents, view their submissions, and perform related academic tasks. This diagram visually captures the different actions each user role can execute within the

system, highlighting the functionality and user interactions essential for effective archiving and management of academic outputs.

User Characteristics

Each of the two users will have different characteristics and different capability with some overlap. The administrator will have an already created account while the student-researchers will need to create an account to access the archive.

The administrator will be able to:

1. Manage user accounts, including listing, viewing details, verifying, and deleting student accounts.
2. Upload files.
3. View uploaded studies.
4. Approve or reject uploaded files, with access to a list of pending files for review.
5. Edit project pages.
6. Access the admin dashboard to modify system information and settings.

The student will be able to:

1. Upload capstone and thesis projects
2. View a list of projects, download, and edit details (if the user is the uploader).
3. Manage account settings for security and personalization.
4. Use the Chatbox to interact with each other and directly message the author of the study.

Functional Decomposition Diagram

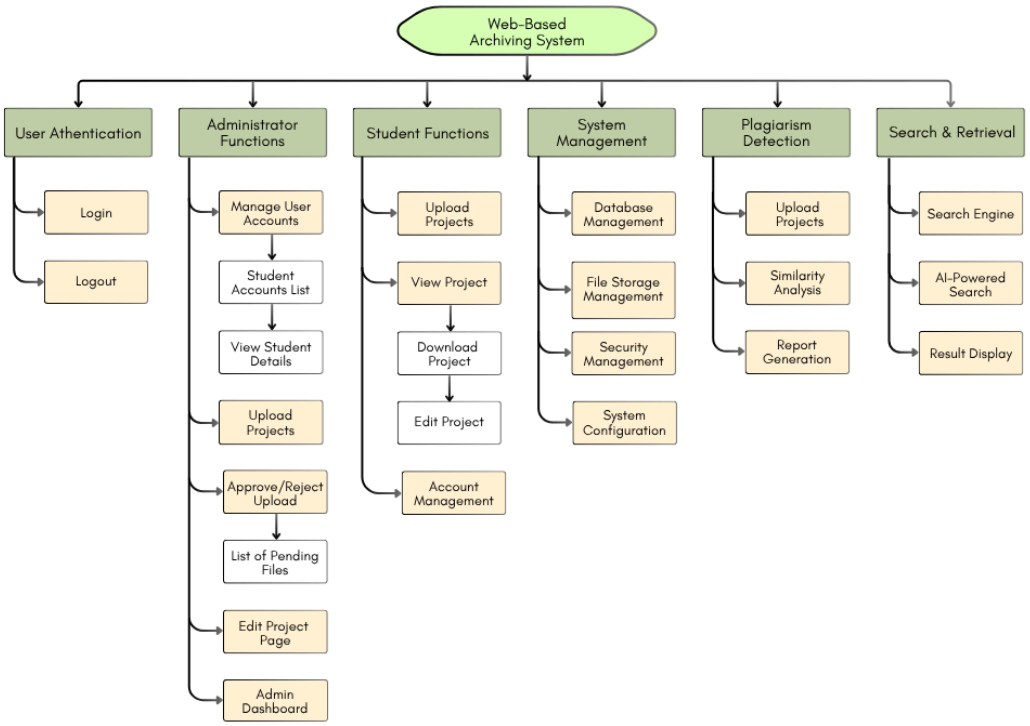

Figure 3. Functional Decomposition Diagram of A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects.

The Functional Decomposition Diagram of the Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects is illustrated above. This diagram outlines the key functions of the system and delineates the various user interactions, providing a clear overview of how the system's components work together to facilitate efficient management and access to archived projects.

System Modules

1. Student Functions Module

This module manages all activities related to student users, including thesis and capstone project submission, viewing submission status, and accessing archived projects. It facilitates students in uploading their work, editing submissions, and managing their personal profiles.

2. System Management Module

Responsible for overseeing the overall operation of the system, this module handles data storage, system configuration, and maintenance tasks. It ensures that system processes run smoothly and manages backups, updates, and integrity checks.

3. Plagiarism Detection Module

This module provides AI-powered plagiarism checking services. It analyzes submitted projects against existing repositories to identify potential instances of plagiarism, generating detailed similarity reports for review by administrators or faculty.

4. Administrator Functions Module

Enables system administrators to manage user accounts, review system logs, perform data backups, and configure system settings. This module ensures the system's security, integrity, and proper functioning.

5. User Authentication Module

Handles login, registration, and user verification processes. It ensures secure access control by authenticating users based on their roles (students, faculty, administrators) and managing session security.

6. Search & Retrieval Module

Provides intelligent search capabilities, allowing users to locate thesis and capstone projects efficiently. Incorporating AI-driven search algorithms, it supports keyword, semantic, and filtered searches to retrieve relevant archived projects quickly.

3.3 DESIGN OF SOFTWARE, SYSTEMS, PRODUCT AND/OR PROCESS

The researchers developed a comprehensive and well-structured design that outlines the inputs, outputs, and procedures. This design aligns with the specified user and system requirements, ensuring compatibility and functional effectiveness throughout the development process.

Conceptual Framework

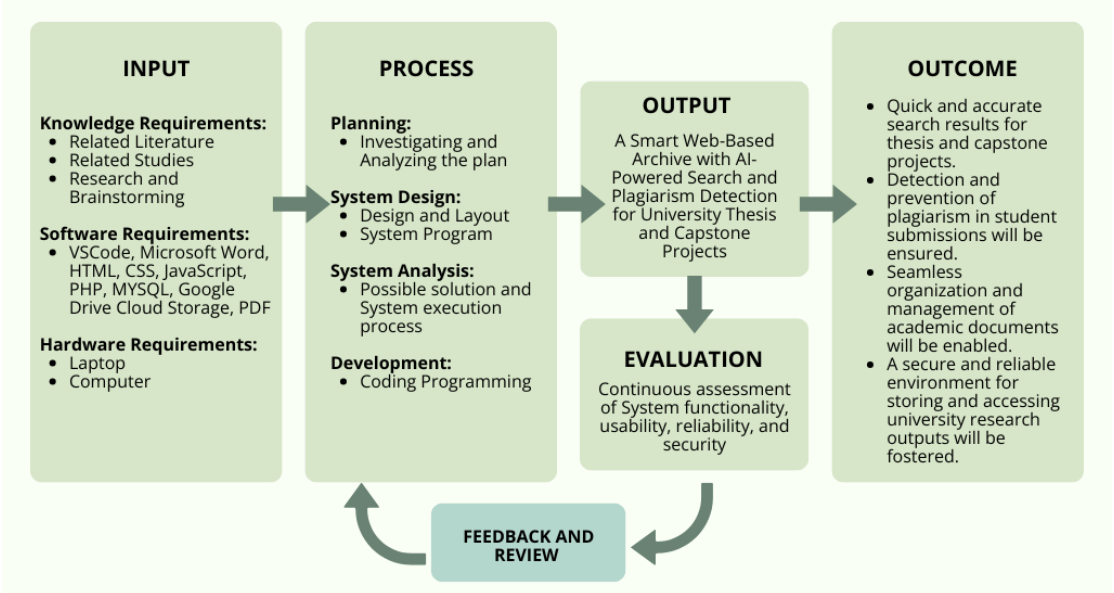

Figure 4. Conceptual Framework of A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects.

Figure 1 illustrates the conceptual framework of the study, utilizing the Input-Process-Output-Outcome-Evaluation model. The Input phase includes related literature and studies, which provide essential background information and context for the research. During the research and brainstorming stages, the researcher gathers

insights and ideas from existing studies, leveraging these to identify techniques and strategies applicable to the project. In the Planning stage, the researcher formulates ideas and outlines the approach to the study, establishing a clear direction. The System Analysis phase involves identifying potential solutions aligned with the study's objectives, examining requirements, and understanding system needs. The System Design phase focuses on developing the website, including visual layout, architecture, and functionality, visualizing how the website will look and operate, serving as a blueprint for development. This framework guides the systematic progression of the research, from initial literature review to the design and development of the proposed system, aiming to achieve a comprehensive and effective solution.

The development phase involves programming activities where researchers create the website. The website will be developed using Visual Studio, which will serve as the primary development environment and programming platform for the system. For database management, SQL Server will be utilized to handle data storage and retrieval. The goal of this study is to develop a Smart Web-Based Archive featuring AI-Powered Search and Plagiarism Detection specifically designed for university thesis and capstone projects.

During the testing phase, researchers will perform systematic testing of each component of the system to identify any issues or feedback requiring modifications. If adjustments are necessary, researchers will revisit the analysis and design phases to incorporate requested changes and ensure system functionality aligns with objectives.

This study aims to provide an efficient, secure, and intelligent platform that simplifies access to academic research outputs, enhances search capabilities, and ensures originality through plagiarism detection, ultimately improving academic resource management.

System Architecture

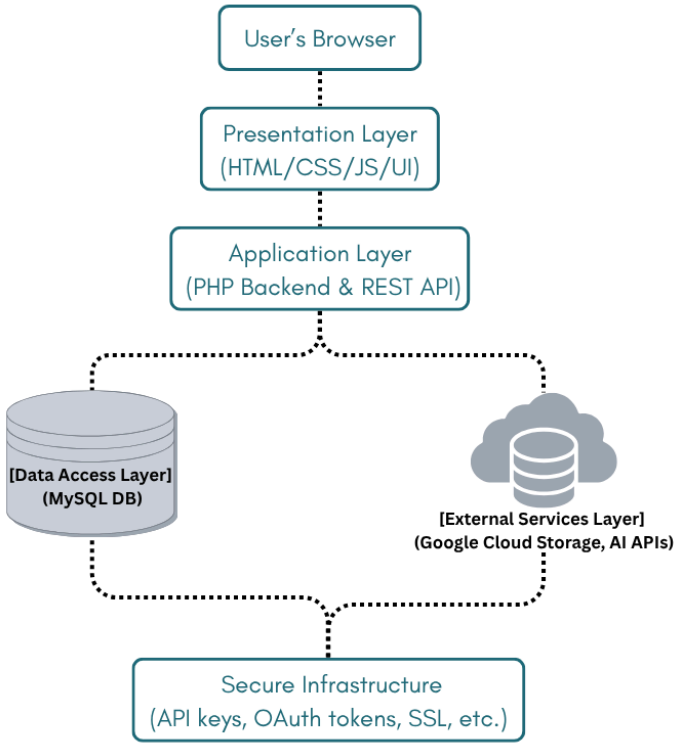

Figure 5. System Architecture of A Smart Web-Based Archive with AI-Powered Search and Plagiarism Detection for University Thesis and Capstone Projects.

This system architecture illustrates a smart web-based archive designed for university thesis and capstone projects. It features a user interface connected to a backend application layer, which interacts with a dedicated Data Access Layer (database) and External Services Layer (AI APIs, cloud storage). Both layers operate within a secure infrastructure that employs SSL, OAuth, and other security measures to ensure data protection, secure communications, and authorized access, enabling efficient search and plagiarism detection for academic submissions.

3.4 DEVELOPMENT AND TESTING

3.4.1 Development Procedure

System Development Methodology

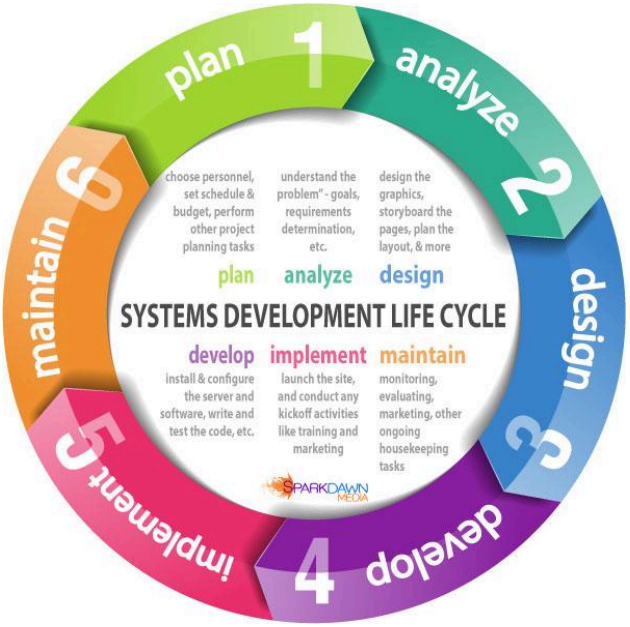

Figure 6. Systems Development Life Cycle

This Figure illustrates the sequential phases of the Systems Development Life Cycle (SDLC), which provides a structured framework for developing and maintaining information systems. This diagram depicts the progression from initial planning through analysis, design, development, testing, implementation, and ongoing maintenance. Each phase is interlinked, ensuring systematic progression and quality assurance throughout the system development process. The SDLC serves as a guide to ensure that the project meets its objectives efficiently, reduces risks, and delivers a functional and reliable system tailored to user needs.

Phase 1. Planning

In this phase, the need for a digital thesis archiving system is identified to replace the manual process. The project scope is defined, and objectives are set to ensure the system effectively manages thesis documents. Feasibility studies are conducted to assess the technical, operational, and economic viability. This stage involves analyzing the current manual archive at the university, identifying gaps, and determining the requirements for a digital solution that streamlines thesis management.

Phase 2. System Analysis

This stage focuses on gathering detailed requirements from stakeholders such as students, faculty, and administrators through interviews and surveys. It aims to understand their needs for search functionalities, user roles, file formats, approval workflows, and security measures. The goal is to create a comprehensive understanding of how the digital archive should function to meet user expectations and safeguard sensitive information.

Phase 3. System Design

In this phase, detailed specifications for the system architecture are developed. This includes designing database schemas for storing thesis information, creating user interface mockups, and outlining workflows for thesis submission, approval, and retrieval. The design ensures that features like admin and user modules, search capabilities, chatbox, and security protocols are properly integrated to facilitate efficient thesis management.

Phase 4. Development

This stage involves building the system based on the design specifications using suitable web development technologies. Key features such as user authentication, thesis upload, approval workflows, search functionalities, and integration with external APIs

like Google API are implemented. The result is a working prototype that demonstrates the core functionalities of the digital archive.

Phase 5. Implementation

This stage involves deploying the system within the university environment. It includes configuring the live server, migrating existing thesis data into the new system, and conducting training sessions for users. The goal is to ensure a smooth transition from manual filing to the digital archive, enabling users to efficiently access and manage thesis documents online.

Phase 6. Maintenance

The final phase provides ongoing support for the digital archive. Regular monitoring ensures optimal system performance, and troubleshooting addresses any issues encountered by users. Future updates, such as adding new features or enhancing security measures, are planned based on user feedback and technological advancements to keep the system up-to-date and reliable.

3.4.2 Testing Procedure

Table 1. Testing Procedures

Component/ Module	Test Conducted	Description	Expected Result
User Authentication System	Login and Registration Testing	Verify users can register, login, and logout correctly.	Users successfully authenticate with valid credentials; invalid attempts are rejected
Thesis Upload Module	Upload Functionality Test	Ensure thesis files can be uploaded with correct metadata.	Files upload successfully; invalid formats are rejected.
Approval Workflow System	Workflow Process Testing	Test the process of thesis submissions,	Theses pass through approval stages

		approvals, and status updates.	correctly; notifications are sent.
AI-Powered Search Functionality	Search Accuracy and Performance Testing	Validate the search results for keywords, author names, and thesis titles.	Relevant results are retrieved accurately and promptly.
Plagiarism Detection Module	Plagiarism Check Accuracy Test	Run sample theses through the plagiarism detection tool.	Plagiarism reports are generated accurately, highlighting similarities.
Security Measures (Encryption, Access Control)	Security Penetration Testing	Test for vulnerabilities such as unauthorized access or data leaks.	System remains secure; unauthorized access is prevented.
External API Integration (Google API)	Integration Functionality Testing	Test the integration with external APIs for location or data retrieval.	External data is fetched and displayed correctly.
User Interface and Navigation	Usability Testing	Check the overall user experience, layout, and ease of navigation.	Users find the system intuitive and easy to navigate.

Table 1.

Testing procedures were guided by industry-standard quality characteristics, including functionality, usability, performance, security, and reliability, to ensure the system meets the required standards for a robust, efficient, and secure web-based archive with AI-powered search and plagiarism detection.

Table 2. Likert Scale

SCALE	RANGE OF MEAN VALUE	INTERPRETATION
5	4.51 – 5.00	Strongly Agree

4	3.51 – 4.50	Agree
3	2.51 – 3.50	Neutral
2	1.51 – 2.50	Disagree
1	1.00 – 1.50	Strongly Disagree

Table 2.

The statistical tool used for data interpretation is the weighted arithmetic mean, as shown in Table 2. The arithmetic mean was employed to determine the average responses for the five-scale options: 5 (Strongly Agree), 4 (Agree), 3 (Neutral), 2 (Disagree), and 1 (Strongly Disagree). The mean scores for each software characteristic were computed to evaluate the overall performance. These individual means were then used to derive an overall evaluation mean, which was interpreted based on the scale provided in the table

3.5 IMPLEMENTATION PLAN

In this section, it describes how the propose system will be deployed, installed and transitioned into an operational system. The plan contains an overview of the system, a brief description of the major tasks involve in the implementation

Table 3. Implementation Plan

Strategy	Activities	Persons Involved	Duration
Approval from School Administration	Letters for the Administrator	Researchers, School Head, Advisers, ICT Coordinator	2 weeks
System Deployment	Deployment on the school's local	Researchers, Advisers, IT	1 week

	server or web hosting platform	Coordinator	
Information Distribution	Announcement through the school's Facebook page, advisory bulletin, and student portal.	Administrator, Researchers, Student	3 Days
Hands-on Training	Lecture and demonstration sessions on utilizing the system.	Researchers, Adviser, Students	2 weeks
Feedback and Evaluation	Gathering suggestions and implementing final adjustments.	Adviser, Students, IT Experts	3 weeks