

Digital Archive Management Information System for Higher Education Institution

Andri Heryandi¹, and Kania Evita Dewi¹

¹ Informatics Engineering, Universitas Komputer Indonesia, Bandung, Indonesia

*Corresponding E-mail: andri@email.unikom.ac.id

Abstract. Universities manage a wide range of documents, including academic, administrative, research, and financial records, which are often scattered across multiple storage locations and formats. This situation leads to inefficiencies in retrieval, risks of duplication, versioning issues, and potential data breaches. To address these challenges, this research analyses, designs, and implements a Digital Archive Management Information System for higher education institutions. The system provides centralized document management with features such as user authentication, file encryption, version control, content indexing, and advanced search capabilities. It also supports interoperability with external storage services like Google Drive, enabling seamless integration of institutional and cloud-based archives. By consolidating scattered documentation into a secure and structured repository, the system improves accessibility, ensures data security, and strengthens institutional readiness for accreditation processes. The result shows that the system contributes to more effective, efficient, and sustainable archive governance, while future enhancements may focus on broader interoperability and intelligent document classification.

Keywords: Digital Archive Management, Information Retrieval, Higher Education, Document Indexing, Data Security, Information System.

1. Introduction

Documentation is the process of creating, managing, storing, and maintaining records or information used to support the operational activities of an institution. A university, as an institution, also engages in documentation processes as part of its operations.

The Digital age has transformed the production and use of documents around the world. In The Information age, the process of entering, delivering, storing, receiving, and categorizing data is crucial (Soltanaga Abbasova, 2020).

The types of documentation typically involved in a university include those related to administrative processes for staff and students (such as diplomas, certificates, and identification documents), academic processes (such as lecture modules, textbooks, assessment sheets, and attendance records), research and community service activities (such as cooperation agreements, grant documentation, research contracts, research reports, and journals), as well as other processes (such as financial reports, decrees, and other legal documents).

One of the obligations that a university must be full fill is the accreditation process. In this process, all documents will be reviewed by assessors. To achieve a favourable accreditation result, a university must ensure a well-organized and comprehensive documentation process.

However, in its efforts to provide an effective documentation process, universities often encounter several challenges, including:

1. Documents are scattered across multiple locations. This occurs because a university typically consists of various faculties, departments, or study programs, each maintaining its own set of documents.

2. Files are stored in multiple locations, such as personal folders on the institution's internal storage (servers), personal storage devices, portable storage media, or cloud-based storage services such as Google Drive. These files may also be stored under different cloud service accounts. This situation makes document retrieval difficult and often leads to duplication of documents.

3. Difficulty in tracking the latest version of a document due to the existence of multiple storage locations.

4. Potential data breaches when storage devices are lost or when files are inadequately protected, allowing confidential information to be easily accessed.

For these reasons, it is essential for a university to have a centralized system that consolidates all document resources in one location, ensuring the institution's operations are well-supported, secure, and efficient.

One way to address this issue is by developing a digital document management information system. As demonstrated in research (Caroline et al., 2022;Lv & Shi, 2020;Wan, 2021), the development of a dynamic information system can help manage documents effectively. However, simplifying the retrieval of well-managed documents presents a new challenge. Therefore, this research will add a document search feature for managed documents. One of the easiest ways to search is by indexing the document, then searching for keywords provided by the user within the document body (Christopher D. Manning et al., 2009).

This research is conducted with the aim of analysing, designing and implementing a Digital Archive Management Information System intended to consolidate scattered documents, improve archive governance, and establish a modern administrative system that is effective, efficient, and sustainable.

2. Literature Review

2.1 Document Management Information System

An information system (IS) can be defined technically as a set of interrelated components that collect, process, store, and distribute information to support decision making and control in an organization (Laudon & Laudon, 2014).

According to (Smallwood, 2013) archive/document/records management is defined as the field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use, and disposition of records, including the processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records.

The digitization process offers many benefits that exceed traditional archival methods. It can preserve invaluable historical data, safeguarding it from natural wear and tear or catastrophic events, such as fire or flood threatening tangible documents (Rowe, 2024).

2.2 Information Retrieval

Information retrieval is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers). According to (Wan, 2021) Information retrieval is concerned with representing, searching, and manipulating large collections of electronic text and other human-language data.

2.3 Indexing

Indexing is the systematic process of analyzing content, extracting key terms or metadata, and organizing them into an index structure to facilitate efficient search, retrieval, and access to the stored information. According to (Büttcher et al., 2010), Indexes provides a mapping between terms and their locations of occurrence in a text collection. Fundamental Components of index be visualize as shown in Figure 1.

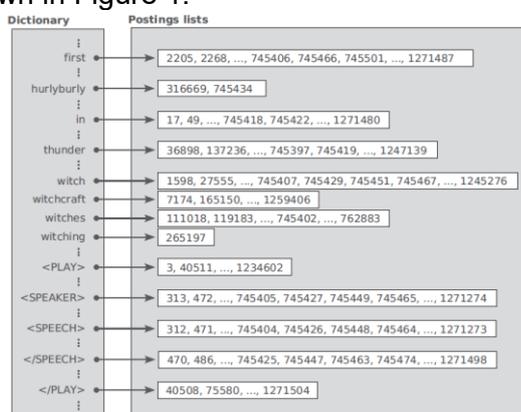


Figure 1. Sample of Fundamental Component of Index

In Figure 1, The dictionary lists the terms contained in the vocabulary of the collection. Each term has associated with it a postings list of the positions in which it appears, consistent with the positional numbering.

To create index, there is process called as determining vocabulary term. This process involve steps are:

1. Tokenization is the task of chopping it up into pieces, maybe at the same time remove certain characters, such as punctuation.
2. Remove stop word.
3. Token Normalization is the process of canonicalizing tokens so that matches occur despite superficial differences in the character sequences of the tokens. This step can include capitalization/case folding.
4. Stemming and lemmatization. Stemming is process that chops off the ends of words and often includes the removal of derivational affixes. Lemmatization is a process to remove inflectional endings only and to return the base or dictionary.

2.4 Encryption

Encryption (Aumasson, 2018) is the principal application of cryptography; it makes data incomprehensible in order to ensure its confidentiality. Encryption uses an algorithm called a cipher and a secret value called the key.

2.5 OpenSSL

OpenSSL (Khlebnikov & Adolfsen, 2022) is an open-source software toolkit that includes a cryptography and SSL/TLS library, as well as command-line utilities that use the library to provide some useful functionality on the command line. The main part of OpenSSL is its library, which means that OpenSSL is mainly useful for software developers.

OpenSSL is also used in web applications. PHP as one of the most used programming languages for web development also has a library that supports OpenSSL operations. These functions are used in PHP to execute OpenSSL functions: `openssl_encrypt` and `openssl_decrypt`.

3. Method

This research uses a Research and Development methodology, since the objective is to design and develop a Digital Archive Management Information System. The research not only investigates the theoretical concepts of digital archive management but also produces a practical system that can be implemented within a higher education institution. The system development stage is using a linear software development process where activities flow sequentially through gathering requirements, design, coding, and testing.

4. Results and Discussion

4.1 System Description

Digital Archive Management Information System is designed as a centralized platform to manage academic and administrative documents within a higher education institution. The system provides a secure and structured environment where documents can be uploaded, indexed, searched, and retrieved efficiently. It incorporates a role-based access control mechanism, where university staff are authorized to manage documents and files, while public users are allowed to view and download permitted content. Archiving aims to provide data and information to those in need as quickly and accurately as possible. To achieve this, effective and efficient archive management is required by understanding the problems within archiving (Faizah et al., 2025).

To enhance data security, all uploaded files are encrypted before storage, ensuring confidentiality and protection against unauthorized access. The system also maintains version control, allowing multiple files to be stored under one document record and enabling retrieval of both the latest and previous versions. An indexing feature is integrated to process file contents, supporting efficient keyword-based search across documents. Additionally, the system allows interoperability with external storage services such as Google Drive, enabling documents stored in the cloud to be downloaded, indexed, and managed within the platform.

By consolidating scattered archives into a single digital repository, the system ensures effective governance of institutional documentation, improves accessibility for stakeholders, and supports accreditation requirements through organized, secure, and sustainable digital archiving.

4.2 System Requirements

Based on the issues described in the background section, the information system to be developed will have the following software requirement specifications:

1. **Login:** The system must provide a user authentication facility. Only users with University Staff status are authorized to perform document management operations.
2. **Download Document:** Public user (including University Staff) may download documents.
3. **Download File:** Public user (including University Staff) may also download files from a specific document, whether in its latest version or in previous versions.
4. **Add File:** University Staff members may upload files to a document. A document may contain more than one file to support multiple document versions. Files uploaded to the server will be encrypted to enhance data security.
5. **Indexing Content:** The system shall provide a facility for indexing file contents to facilitate the search process.
6. **Searching:** The system shall feature document search capabilities based on title, and file content.
7. **Download File from Google Drive:** The system shall provide a facility to download files from the cloud server for indexing purposes.
8. **Delete File:** Document management shall include a facility for deleting files from a document.
9. **Access Management:** The system shall include a feature to set access permissions for documents/archives.
10. **Document Management:** The system shall provide a document management facility.

4.3 Use Case Diagram

A use case diagram is one of the diagrams in Unified Modeling Language (UML) used to model the functional requirements of a system. Use case diagram visually shows the interactions between actors (users or external systems) and the use cases (the system's functions or services).

The use case diagram for this system can be seen in Figure 2. The actors involved in this system are:

1. **Public:** Any individual who can access the system (students, employees, lecturers, administrators, or the public).
2. **University Staff:** Users authorized to manage documents/files (Faculty Staff, Department Staff, Lecturer).
3. **Google Drive:** An external file storage service used outside the institution's internal storage infrastructure.

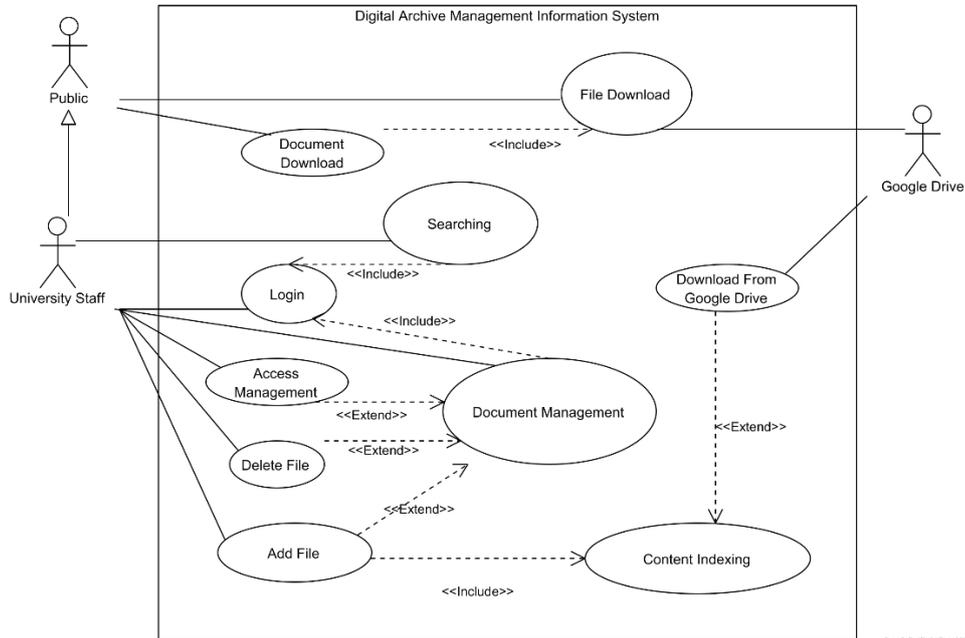


Figure 2. Use Case (Digital Archive Management Information System.png)

4.4 Activity Diagram

An Activity Diagram is a type of diagram that models the workflow of activities and actions in a system or process.

Figure 3 shows the activity diagram for Document Management Use Case.

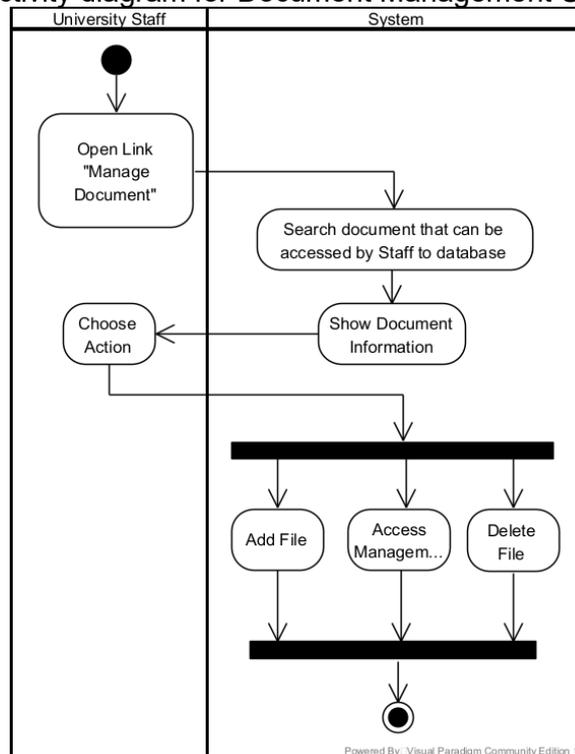


Figure 3. Document Management Activity Diagram

Figure 4. shows the activity diagram for Add File use case

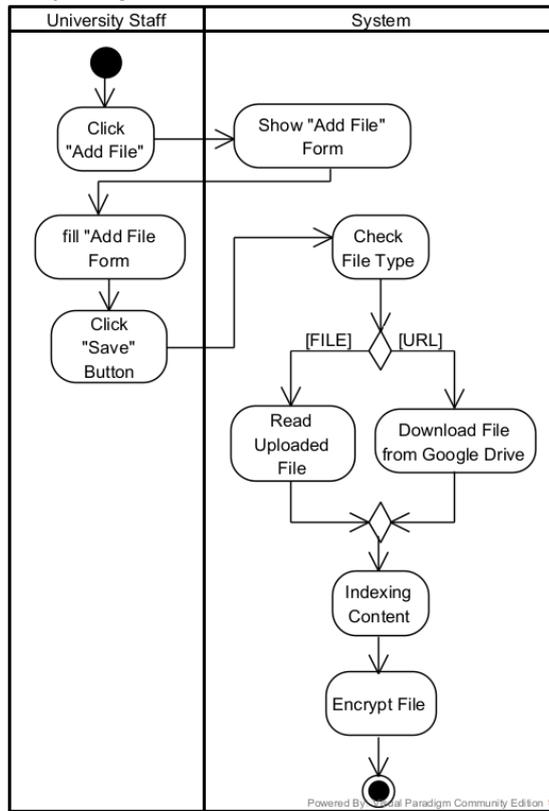


Figure 4. Content Indexing Activity Diagram

Figure 5 shows the activity diagram for Searching use case

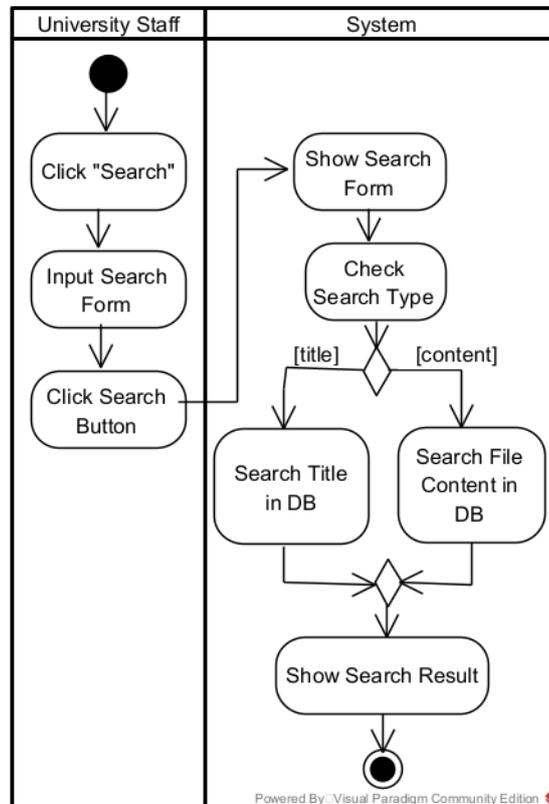


Figure 5. Searching Activity Diagram

4.5 Database Design

A database schema is the blueprint or structure of a database that defines how data is organized, stored, and related to each other. It describes the tables, fields, relationships, constraints, and other elements that make up the database.

Based on the requirements, this system needs to save document information in a database. Figure 6 show the database schema for the system.

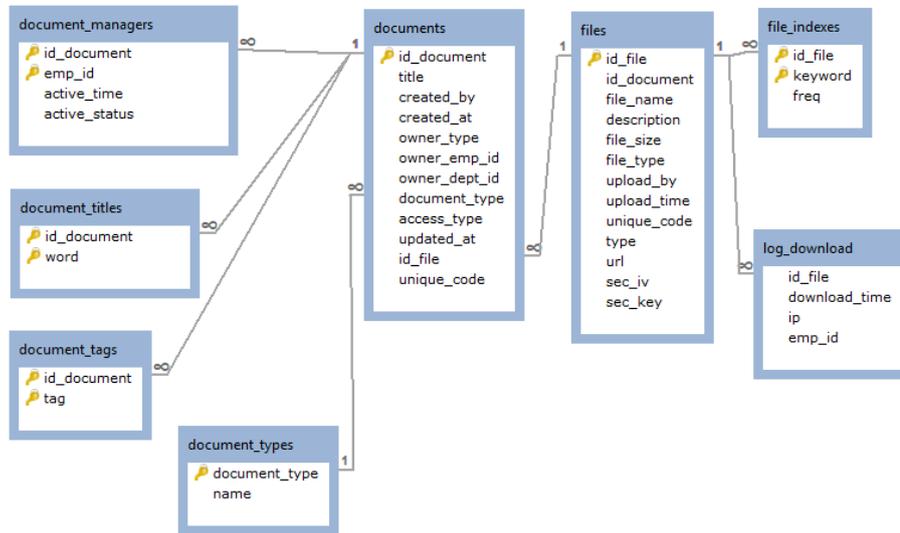


Figure 6. Database Schema

4.6 Implementation

Figure 7 shows the screenshot of Document Management. In this screen, a user can add new document, copy URL of document to clipboard, download document and setting document information.

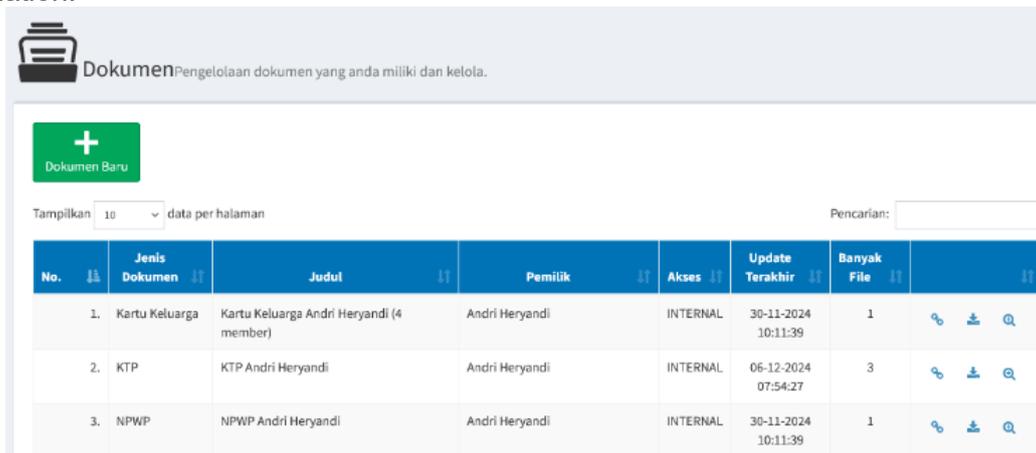


Figure 7. Document Management Interface

Figure 8 shows the screenshot of Add File activity. In this screen, user can input the information about the document such as document type, title, tag, owner type and the owner name of the document, the access mode, the file location and description of document.

Penambahan Dokumen Baru

Jenis Dokumen* Akta Kelahiran

Judul* Akte Kelahiran Andri Heryandi

Tag akte kelahiran andri heryandi

Pemilik Dokumen* Pribadi (Andri Heryandi)

Nama Pemilik* Pilih Institusi Pemilik

Akses* Public - Semua

File 1994-xx-xx - Akte Kelahiran.png
Satu file per upload. Boleh dikosongkan, file bisa diupload nanti. Konten file docx dan pdf akan diindex untuk mempermudah pencarian.

Keterangan File Keterangan/Deskripsi File

Figure 8. Add File interface

Figure 9 shows the screenshot of Searching. In this screen, user can find document by title of document or tag of document or text in the content of the document file.

Pencarian Mencari dokumen yang boleh anda akses.

Pencarian di Judul Dokumen

Hasil Pencarian

Tampilkan data per halaman

No.	Judul Dokumen / Nama File	Pemilik	Pembuat	Waktu Pembuatan	Terakhir Update	Ukuran File	
1.	Akte Kelahiran Deny Angga Diredja	Deny Angga Diredja	Deny Angga Diredja	30-11-2024 09:33:54	30-11-2024 09:33:55	1.084 KB	<input type="button" value="Cari"/>
2.	Akte Kelahiran Agus Mulyana	Agus Mulyana	Agus Mulyana	30-11-2024 09:37:29	30-11-2024 09:37:29	1.797 KB	<input type="button" value="Cari"/>

Figure 9. Searching Interface

5. Conclusion

This research has successfully developed a Digital Archive Management Information System for higher education institutions that addresses challenges such as scattered document storage, version control, and data security. The system integrates key features including user authentication, file encryption, indexing, and advanced search to ensure efficient, accessible, and secure document management. Its centralized approach not only supports institutional operations but also strengthens accreditation readiness, while future improvements may enhance interoperability, cloud integration, and intelligent search capabilities.

References

- Aumasson, J.-P. (2018). *Serious Cryptography*. William Pollock.
- Büttcher, Stefan., Clarke, C. L. A. ., & Cormack, G. V. . (2010). *Information retrieval : implementing and evaluating search engines*. The MIT Press.

- Caroline, D. A., Ismanto, B., & Rina, L. (2022). Implementation of digital archives using a dynamic archive information system. *Jurnal Kajian Informasi & Perpustakaan*, 10(2). <https://doi.org/10.24198/jkip.v10i2.33203>
- Christopher D. Manning, Prabhakar Raghavan, & Hinrich Schütze. (2009). *An Introduction to information retrieval*. Cambridge University Press.
- Faizah, N. A., Thohir, M., Salem, S., & Mardiyah. (2025). Optimizing Digital Archive Management to Improve the Quality of Integrated Public Services. *Kharisma: Jurnal Administrasi Dan Manajemen Pendidikan*, 4(1), 31–43. <https://doi.org/10.59373/kharisma.v4i1.64>
- Khlebnikov, Alexei., & Adolfsen, Jarle. (2022). *Demystifying cryptography with OpenSSL 3 : discover the best techniques to enhance your network security with OpenSSL 3.0*. Packt Publishing.
- Laudon, K. C. ., & Laudon, J. P. . (2014). *Management information systems : managing the digital firm*. Pearson Education.
- Lv, Z., & Shi, H. (2020). The Exploring on University Archives Management System Based on Information System. *Journal of Physics: Conference Series*, 1550(3). <https://doi.org/10.1088/1742-6596/1550/3/032017>
- Rowe, R. (2024). Digital Records Management: The History of Your City. In *Certified Public Manager® Applied Research* (Vol. 5, Issue 1). <https://scholarworks.sfasu.edu/cpmar>.<https://scholarworks.sfasu.edu/cpmar/vol5/iss1/1>
- Smallwood, R. F. (2013). *Managing Electronic Records : Methods, Best Practices, and Technologies*. John Wiley & Sons, Inc.
- Soltanaga Abbasova, V. (2020). *Main Concepts of the Document Management System Required for Its Implementation in Enterprises*. 1, 2020. <https://doi.org/10.21303/2313-8416.2020.001149>
- Wan, L. (2021). Digital web archive collection framework based on intelligent management information system and optimization. *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, 1–4. <https://doi.org/10.1109/ICICCS51141.2021.9432158>