

Centralized Document Management System with Secure Cloud Storage - DocuFlow

Sonali Bihari Buwade

Department of Science and Technology, G.H.Raisoni Skill Tech University, Nagpur, India

ABSTRACT

DocuFlow, another name for a centralized document management system with secure cloud storage (DMS), is an web app that keeps all types of documents in one location. Employees and admin are among the access-based roles that are used by the DocuFlow management system, which keeps the document in cloud storage. The web application is protected by encryption and authentication. Everything is safely kept in the cloud and controlled via a single platform that is backed up on cloud storage, as opposed to being stored on local devices or dispersed folders. Some issues with the traditional file management system include lost files, improper access control, version control, and security. It is also difficult to determine who made changes without a central backup.

Based on the given problems of storing documents in traditional ways, a centralized document system with a secure cloud solution comes up. A centralized DMS solves all of these problems, and many Industries using DMS in the following sectors: banks, Hospitals, Colleges, IT companies, and Government Offices. DMS provides centralized storage for documents, ensuring secure access using authentication and authorization with role-based access control (RBAC)

While maintaining document version history and activity logs with encrypted cloud storage, enable easy retrieval and search.

Keywords: Document System; Cloud Based Document System; DocuFlow; DocuFlow Document System.



This is an open-access article under the [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

1. Introduction

In today's digital world, organizations generate a tremendous volume of digital documents on a daily basis. Documents are frequently stored across numerous systems, local drives, emails, and physical storage, resulting in duplication, data loss, inadequate version control, and security issues. In addition, the workplace culture is changing. Remote employment is gaining popularity, which increase the dependence on digital documents instead of paper-based system or traditional document management (TDM), So that an employee can access documents without any limitation. However, in traditional document management (TDM), documents are held in physical form with limited access control, which increases data security and policy violations. Furthermore, the TDM has problems tracking document changes, weak backup and recovery procedures, manual processing resulting in slowness, and a significant danger of unauthorized access.

There are a number of security risks associated with the documents uploaded to digital storage, including the possibility of unauthorized access to sensitive organizational data, the possibility of data

breaches and cyber threats, the possibility of insiders abusing the documents' data, the absence of encryption and access control in basic storage systems, and the lack of appropriate audit tracking mechanisms. The goal of the document management system is to safeguard important organizational data, not just save files. A comprehensive infrastructure that offers processing power, storage, security, and availability via the internet, cloud technology is more than just online storage. Cloud technology is essential to the scalability, security, accessibility, and dependability of a centralized document management system. If storage becomes full, new hardware must be installed, configured, and maintained manually. Traditional storage systems require physical servers. Cloud platforms that offer on-demand scalability include Amazon Web Services and others, which increase capacity automatically as needed; no hardware needs to be installed for more storage; businesses only pay for what they use; and cloud storage's high availability ensures that the system is always available. When one server fails, another server takes over. Cloud infrastructure greatly lowers the danger of document management system outages, which can cause delays in approvals, interfere with workflows, and impact document management systems. Cloud storage helped with geo-replication (storing data in many locations), data redundancy, and automatic backups by lowering infrastructure costs. Documents can be accessed from anywhere, and if one site fails, they can be recovered from another.

With increase growth of remote and hybrid work environments, multiple employees across the organization accessing the same documents, these documents can be accessed using role-based system, which helps with the misuse of document's data and is important for maintaining document integrity and version control. Centralization is necessary because it provides a single platform for organizing, storing, and retrieving documents, reducing duplication and redundancy while enhancing accountability, transparency, and administrative control.

1.1 Motivation

The creation and storage of digital documents, including contracts, financial records, personnel data, and administrative reports, has increased as a result of the quick digitization of organizational operations. Many businesses still use storage technologies like local drives and unprotected shared folders in spite of this expansion. These acts lead to data duplication, inefficient retrieval, and a higher risk of information loss because they lack centralized rules, structured access control, and appropriate version control. Secure multi-employee access to shared documents is becoming more and more necessary.

1.2 Contribution

The following is a list of the paper's key contributions:

1. To design and develop Centralized DMS with secure cloud storage.
2. To apply role-based access control for document system.
3. To integrate secure authentication and authorization mechanisms to prevent unauthorized access in DMS.
4. To incorporate encryption techniques to ensure data security during storage and transmission.
5. To implement document version control and audit logging mechanisms for maintaining integrity and accountability.
6. To evaluate the system in terms of security, scalability, and performance in a cloud environment.

The paper is organized as follows. DMS's development status and my contributions are briefly discussed in Section 1, and related efforts are discussed in Section 2. In Section 3, I describe the architecture, innovation, and benefits of our document management approach in great detail; in Section 4, we explain our experimental setup and provide results; in Section 5, we summarize and introduce future work.

2. Related work

Many studies have been conducted in the field of document management systems with secure cloud storage in the past years. Researchers have concentrated on creating systems that improve document accessibility while preserving data security and integrity as more businesses move toward a digital world. Several cloud-based document storage frameworks have been proposed to increase scalability, cost effectiveness, and availability.

Some studies have discovered cloud-integrated document management architecture using platforms such as Amazon Web Services, Microsoft Azure and many more others, these systems emphasize centralized storage and distributed accessibility but mostly rely heavily on default cloud security config.

While these approaches provide scalability and high availability, but they still lack the fine-grained access control mechanisms and structured document governance policies, leading to potential security vulnerabilities. Other research efforts have focused specifically role-based access Control models in DMS. These studies highlight the importance of limited access control based on the roles to reduce insider threats and unauthorized access and modifications.

The encryption-based document protection methods have also been widely studied and then later the researchers introduced the methods of encryption techniques to secure data both at rest and in transit. Many existing research provides valuable contributions in the areas of cloud scalability, access control and encryption there still a need for a unified framework that can integrated centralized management, secure authentication, encryption policies, role-based authorization, version control. This research aims to address this gap by proposing a centralized DMS that combines theses essential security and rules features into structured platform.

3. Research Methodology

3.1 Problem statement

Today, all around the world, many organizations generate and manage large volumes of digital documents, which include contracts, reports, academic records, financial data, and admin files. However, these documents are often stored in many different locations, such as local systems, email servers, external drives, and many more. This fragmented storage approach results in poor version control, ineffective document retrieval, and difficulty in monitoring user access. As a result, organizations have less productivity and more operational complexity.

In addition to inefficiency, the absence of centralized governance and structured access control mechanisms exposes sensitive information to security risks, which includes unauthorized access, data breaches, and insider misuse. Existing system lacks proper encryption, role-based access control, audit tracking and secure backup mechanisms. With the rise in demand for remote collaboration and cloud accessibility, there is a critical need for centralized, secure and scalable document management system (DMS) that ensure data confidentiality, integrity, availability, and efficient collaboration within organization.

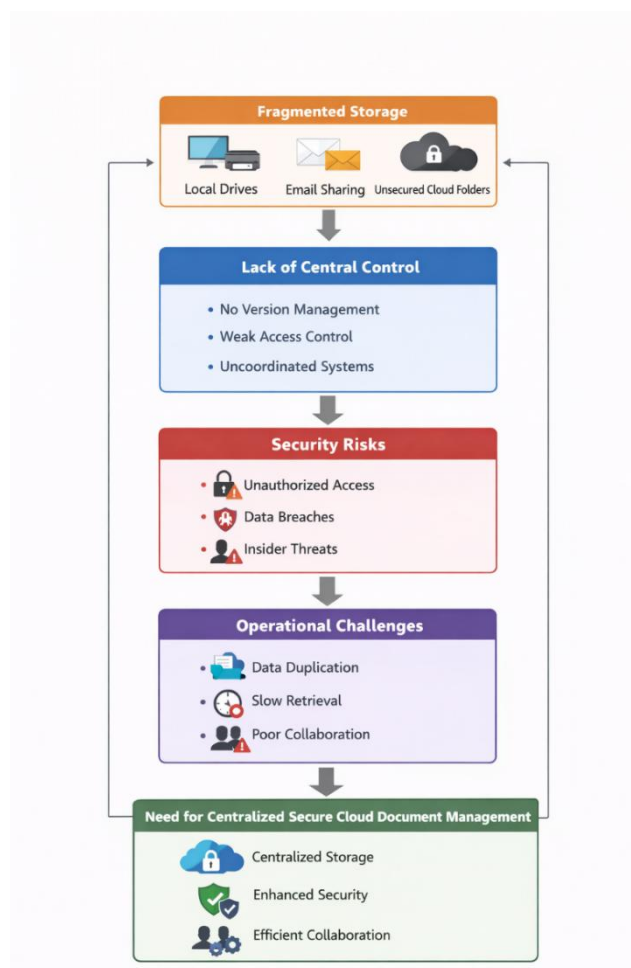


Fig 1. Block diagram of the problem statement

3.2 Decentralized Document Storage Environment

In modern organization, digital documents contracts, reports, invoices, academic records and admin files are generated daily. These documents usually stored across multiple places which includes local systems, email systems, external drives etc. Due to this decentralized storage structure, documents become scattered across different locations without any control.

3.3 Lack of Version Control and Document Tracking

In TDM system, there was no proper structured version management mechanism. When multiple users edit or share the same file, different versions are created without proper synchronization.

This leads to:

1. Confusion between documents versions.
2. Overwriting important data in the documents.
3. Loss of documents history.
4. Inefficient to trace the modifications within documents.

The absence of audit log and tracking mechanisms makes it difficult to monitor document lifecycle and employee's activity.

3.4 Weak Access Control and Security Mechanisms

Most already existing document storage approaches simply rely on basic authentication methods without role-based access control. Documents are frequently shared through unsecured links.

The major security concerns include:

1. Unauthorized access to confidential documents.
2. Data breaches and cyber threats.
3. Insider misuse of sensitive information.
4. Lack of encryption and secure backup systems.

Without centralized DMS it is challenging to maintain data confidentiality, integrity and security.

3.1.4 Inefficient Collaboration and Retrieval

In shatter document environment, employees spend extravagant time on searching files across different systems. Document retrieval becomes difficult and inefficient due to lack of indexing and structured organization.

Challenges which include:

1. Time taken will to less productivity.
2. Reduce team collaboration.
3. Increase admin workload.

These limitations reduce overall productivity and efficiency.

3.1.5 Centralized Secure Cloud -Based DMS - needs

Considering all the limitations of decentralized document handling, there is a strong need for a centralized, secure and scalable Document Management Systems (DMS) with cloud storage.

The system must include:

1. Centralized document repo.
2. Role-based access control system.
3. Version management in DMS.
4. Data encryption and secure backup.

4. Methodology

4.1 Research Methodology

The research methodology describes the systematic approach that is used to design and develop the proposed centralized DMS with the secure cloud storage. The main focus of this methodology is to create a secure, efficient and scalable webapp that enables organizations to store, manage, and share documents from a centralized location while maintaining strict security protocols.

The prefer system integrates cloud-based storage, secure authentication mechanisms, and structured document management techniques to ensure efficient document handling. The main focus of this methodology is on several key aspects including system architecture design, document upload and storage mechanisms, document upload and storage mechanisms, access control implementation, data security measures and document retrieval functionality. Modern web technologies and cloud computing services are utilized to build a reliable and scalable document management environment.

The research process begins with analyzing the limitation of existing document storage methods and identifying, the need for centralized management. Based on this research, a system architecture is designed that integrates a web-based interface, backend server, DMS, and secure cloud storage infrastructure. The system enables users to upload documents, store metadata, control access permissions, and retrieve files efficiently through a centralized platform.

Furthermore, security mechanisms such as authentication, role-based access control, encryption, and activity logging are incorporated to ensure data confidentiality and integrity. The developed system is evaluated based on performance, usability, security and scalability to determine its effectiveness in managing organizational documents securely in a cloud environment.

Through this structured methodology, the prepared system provides aims to provide an effective solution for managing large volumes of digital documents while supporting secure collaboration and efficient document lifecycle management.

4.1.1 System Architecture

The prepared centralized DMS with secure cloud storage is designed using a multi-layer 3-tier architecture which improves security and scalability that is integrated with the user interface (Frontend), application server (Backend) and with the Database and additionally using cloud storage for cloud storage services. This 3-tier architecture ensures the documents efficient management.

The system consists of 3 main components/parts: The Frontend, Backend and Databases. The employee/user interact with the frontend web application where they can upload, download, manage and share documents. The backend server processes these requests coming from frontend, manages authentication and authorization, and handles communication between the database and cloud storage. The database stores the uploaded documents metadata information such as file name, file type, upload date and user access permission so the database can manage these files effectively in the database without any duplicate or overwriting problems.

The cloud storage is used to store the original document securely, allowing employee/user to access files from different locations while maintaining centralized control over these. The architecture ensures that documents are stored securely while maintaining efficient retrieval and collaboration among authorized employee/user.

The system also provides efficient search and retrieval functionality. Users can easily locate documents using metadata information such as document name, file type, upload date, or keywords. This feature significantly reduces the time required to find important documents within large datasets and improves overall organizational productivity.

In addition, the centralized architecture simplifies system maintenance, monitoring, and data backup. Regular backups of stored documents ensure that data can be recovered in case of accidental deletion, system failure, or cyber threats. By combining centralized management with cloud-based storage, the proposed system offers a reliable, scalable, and secure platform for managing organizational documents efficiently.

4.1.2 Document Upload and Storage Mechanism

The document upload mechanism allows employee/user to upload a document (in formats pdf, txt, word, etc.) in centralized system through a web interface. When employee/user uploads a document, the system first verifies the employee/user on the role-based authentication and authorization right before allowing the upload process.

Once the document gets uploaded through web interface the documents get stored in the cloud storage and the documents metadata information stores in the database. The document contains metadata information's like filename, original filename, file size, uploaded by, timestamp, mimeType, cloudUrl, description, tags, category. This separation between document storage and metadata storage improves system efficiency and enables faster documents retrieval.

This system also supports version control, giving employee/user to update document without losing original document's version. This feature ensures that the document history is well maintained and prevent accidental loss of important information.

Columns

- ✓ Filename – filename of the document.
- ✓ Original name – original name of the document uploaded by employee/client.
- ✓ File size – size of the file in bytes.
- ✓ mimeType – Type of the document.
- ✓ Description – Extra information about the document and this helps in document organization.
- ✓ Tags – keywords related to the document.
- ✓ Category – Used to group documents, this helps with filtering documents in the system.
- ✓ Uploaded by – Stores who uploaded the file.

4.1.3 Access Control and Authentication

Security is the main important component of the proposed document management system (DMS). The system implements authentication mechanisms to verify the employee's identity on these mechanisms before granting any access to them.

Role-based Access Control (RBAC) is used to grant document access permissions, based on different role such as admin and employee, each role has specific permissions that determine what actions a individuals can perform within the system, such as uploading, downloading, deleting, updating, assigning or viewing restricted documents.

The structured access control mechanism ensures that the restricted document only access by authorized individuals only, these prevent unauthorized access and maintain its data security.

The RBAC mechanism ensures that sensitive documents are protected by restricting access based on user roles and responsibilities. Only authorized individuals are allowed to access restricted documents, thereby reducing the risk of unauthorized data access or misuse. This structured access control approach improves data confidentiality and ensures that users can only interact with documents relevant to their job responsibilities.

In addition to role-based permissions, the system also maintains activity logs to track user actions within the platform. These logs record operations such as document uploads, downloads, edits, and deletions, allowing administrators to monitor system activities and detect any suspicious behaviour. This auditing capability enhances system transparency and accountability while providing an additional layer of security.

Furthermore, secure communication protocols are used to protect data transmission between the user interface and the backend server. By combining authentication, role-based access control, and activity monitoring, the proposed system provides a robust security framework that protects organizational documents from unauthorized access, data breaches, and potential security threats.

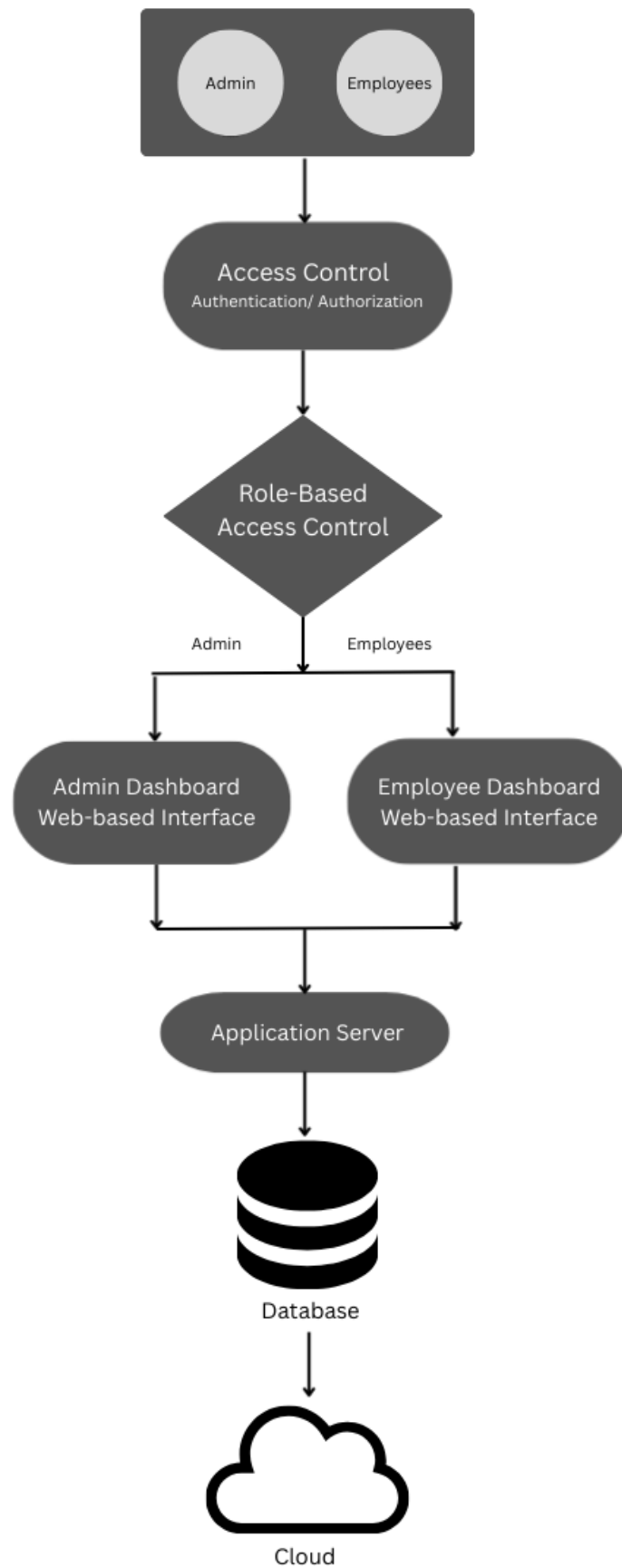


Fig 3. Flow chart of Access Control and Authentication

4.1.4 Data Security and Encryption

The Data Security and Encryption ensure the confidentiality and integrity of stored documents; the proposed system incorporates strong security mechanisms which includes encryption and secure communication protocols.

Documents are protected using encryption in cloud storage, this ensure that even if unauthorized access occurs, the data still remains unreadable without proper decryption keys, also secure communication protocols such as HTTPS are used to protect data transmission between the client interface and the server.

4.1.5 Document Retrieval and Search Mechanism

Effective document retrieval is an important feature of the Document Management System (DMS). The DMS also includes search, category, filtering mechanism that allows employees to quickly locate documents based on their metadata information such as filename, originalname, mimeType, file size, description, tags, upload by and timestamp.

The search functionality improves productivity by reducing the search time required to find documents. Additionally, employees can also download, read, update and view directly from the centralized platform depending on their access permissions.

By combining storage, metadata indexing, and efficient search algorithms, the system ensures quick and reliable document retrieval for authorized employees.

5. Conclusion and Future Work

The Centralized Document Management System with Secure Cloud Storage provides an efficient solution for managing and storing document in a secure cloud storage in a organized manner. In modern organizations, the growth of digital data has been created numbers of difficult challenges in storing, managing, and accessing documents. Traditional documents handling methods often lead to issues such as data redundancy, version control, unauthorized access. The proposed system considers these challenges by providing a centralized platform where documents can be securely stored, accessed, and managed through cloud-based infrastructure.

The system integrates important features such as role-based access control, authentication mechanisms, encrypted cloud storage, document version management and secure sharing capabilities. These features only authorized users can access specific documents while maintaining the confidentiality, integrity and availability of data. Additionally, the centralized structure improves document retrieval efficiency and enhances collaboration among employee by allowing secure document sharing and real-time access. Overall, the proposed system improves organizational productivity, strengthens data security, and provides a scalable framework for efficient document management in cloud environments.

Although the proposed Centralized Document Management System provides a secure and efficient solution for document handling, several important can be implemented in the future to enhance its functionality and performance. One major potential improvement is to integrate advanced artificial intelligence technique for automatic document classification, content extraction, and smart search capabilities. This would allow the system to automatically organize documents based on their content and improve document retrieval efficiency.

Another possible enhancement is the implementation of blockchain technology to strengthen document security and maintain a tamper-proof record of document access and modifications. Real-time collaboration can also be introduced in the future work where it allows multiple employees can edit and manage documents simultaneously. Furthermore, mobile application supports and advance analytics dashboard can be introduced to improve accessibility and provide better insights into

document usage patterns. This improvement will further enhance the scalability, security, and usability of the centralized document management system in modern digital environments.

5. Reference

1. Michael Armbrust et al., "A View of Cloud Computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
2. Peter Mell and Timothy Grance, "The NIST Definition of Cloud Computing," NIST Special Publication 800-145, 2011.
3. Ian Foster et al., "Cloud Computing and Grid Computing 360-Degree Compared," *Grid Computing Environments Workshop*, 2008.
4. Ravi Sandhu et al., "Role-Based Access Control Models," *IEEE Computer*, vol. 29, no. 2, pp. 38–47, 1996.
5. William Stallings, *Cryptography and Network Security: Principles and Practice*, Pearson, 2017.
6. Thomas Erl, *Cloud Computing: Concepts, Technology and Architecture*, Pearson, 2013.
7. Rajkumar Buyya et al., *Mastering Cloud Computing: Foundations and Applications Programming*, Morgan Kaufmann, 2013.
8. Kai Hwang, Geoffrey Fox, and Jack Dongarra, *Distributed and Cloud Computing*, Morgan Kaufmann, 2012.
9. Andrew S. Tanenbaum and Maarten Van Steen, *Distributed Systems: Principles and Paradigms*, Pearson, 2017.
10. Nina Godbole and Sunit Belapure, *Cyber Security: Understanding Cyber Crimes*, Wiley, 2011.
11. John R. Vacca, *Cloud Computing Security*, CRC Press, 2014.
12. Dimitrios Zissis and Dimitrios Lekkas, "Addressing Cloud Computing Security Issues," *Future Generation Computer Systems*, 2012.
13. Mladen A. Vouk, "Cloud Computing – Issues, Research and Implementations," *Journal of Computing and Information Technology*, 2008.
14. Craig Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," *STOC Conference*, 2009.
15. Abhishek Dubey et al., "Secure Cloud Storage System Using Encryption," *International Journal of Computer Applications*, 2015.
16. International Organization for Standardization, "Information Security Management Systems – Requirements," ISO/IEC 27001, 2013.
17. National Institute of Standards and Technology, "Guide to Cloud Computing Security," NIST Publication 800-144, 2011.
18. Chris Reed, "Information Governance and Document Management Systems," *Information Management Journal*, 2016.
19. Rittenhouse John W. and James F. Ransome, *Cloud Computing: Implementation, Management, and Security*, CRC Press, 2017.
20. Eric Yuan, "Cloud Collaboration Systems for Enterprise Applications," *IEEE Cloud Computing*, 2018.
21. K. H. Pradeep, "Secure Document Sharing Using Cloud Storage," *International Journal of Computer Science and Information Technologies*, 2019.

22. Saurabh Singh et al., “A Survey on Cloud Computing Security Issues and Challenges,” *Journal of Network and Computer Applications*, 2018.
23. Florian Skopik, *Security and Privacy in Cloud Computing*, Springer, 2014.
24. Khaled Salah et al., “Blockchain-Based Secure Document Sharing System,” *IEEE Access*, 2019.
25. Nabil Sultan, “Cloud Computing for Education: A New Dawn,” *International Journal of Information Management*, 2010.
26. M. AlZain et al., “Cloud Computing Security: From Single to Multi-Clouds,” *HICSS Conference*, 2012.
27. Siani Pearson, “Privacy, Security and Trust in Cloud Computing,” *Springer Cloud Computing Series*, 2013
28. Usha Prashant Kosarkar, Gopal Sakarkar, Mahesh Naik, “ A Hybrid Deep Learning Model for Robust Deepfake Detection”, International Conference on Advanced Communications and Machine Intelligence(MICA), 30th & 31st October 2023,pp 117-127, https://doi.org/10.1007/978-981-97-6222-4_9