JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SECURE CONTENT BASED IMAGE RETRIEVAL AND COPY DETERRENCE USING INVISIBLE DIGITAL WATERMARKING

[1]Adhila Fathima A S[1], [2]Thasneem H[2] , [3]Mrs.S. Devilakshmi[3]
Department of Information Technology, Meenakshi Engineering College , Tamil Nadu
(Affiliated to Anna University Tamil Nadu)

Abstract :
Ensuring the security and integrity of data is crucial in the current digital environment. With a focus on collaboration between data owners, users, and administrators and the integration of cutting-edge methods like encryption and invisible digital watermarking, this project offers a comprehensive solution for data management. The three primary components of the system are administrators, users, and data owners. Owners of data have the ability to upload a variety of assets, such as documents, photos, and PDFs, and add undetectable digital watermarks on them. These watermarks are used for tracking and ownership verification since they contain encrypted text. After authenticating, users get access to a personalized dashboard where they may browse the information of files that data owners have submitted. Users receive an email containing a public-private key combination after choosing a file to download. The file and its concealed watermark can be safely retrieved and viewed with this key pair. A key component in guaranteeing data security across the system is encryption. Sensitive data, such as user information, file details, and key information, is encrypted and decrypted using AES encryption and stored in the database. Administrators can monitor key distributions, file interactions, and user activity with privileged access. They are able to keep an eye on file downloads, follow user activities, and make sure security protocols are followed. The architecture of the system places a high priority on data privacy and integrity, protecting against illegal access and tampering with strong encryption techniques and undetectable digital watermarking. The project offers a strong foundation for safe data management and cooperation in digital environments by combining these cutting-edge technologies.

Index Terms **– Data owners, users and administrators, Encryption, Invisible Digital Watermarking, Public-Private key, AES encryption, Database, Data Privacy, Integrity.**

I. INTRODUCTION

The management and security of data have emerged as critical issues for both individuals and enterprises in today's linked digital environment. The rising amount of private data being shared and kept online has made it more difficult to guarantee the confidentiality, integrity, and validity of data. This project offers a comprehensive solution for safe data management in response to these issues by utilizing cutting-edge methods including encryption and undetectable digital watermarking. The cooperation of data owners, users, and administrators inside a safe and reliable framework is the foundation of this solution. Data owners, who are responsible for uploading different kinds of information, such as papers, photos, and PDFs, have the ability to add invisible digital watermarks to their assets. These encrypted text watermarks are useful for tracing and monitoring file interactions in addition to being a reliable method of confirming file ownership. Users are given an easy-to-use interface via which they may browse and download files submitted by data owners after logging in to the system. Users receive a distinct public-private key pair via email, adding extra protection to the file-downloading process. This key combination not only makes sure that the files can be viewed and recovered safely, but it also makes it easier to decrypt any hidden watermarks that may be present. A key component in protecting sensitive data across the system is encryption. User passwords and other sensitive data are encrypted and decrypted in the database using the Advanced Encryption Standard (AES).

The project guarantees strong security against illegal access and data breaches by utilizing AES encryption.With privileged access, administrators are essential to monitoring system performance and implementing security measures. They are able to keep an eye on user behavior, record file exchanges, and make sure that rules are followed. Administrators protect the security and

integrity of data in the system by monitoring it. Through the integration of cutting-edge technologies like AES encryption and undetectable digital watermarking, this project seeks to offer a dependable and safe platform for data management and communication in digital contexts. The initiative tackles the changing difficulties of protecting sensitive data in a connected world by taking a holistic approach to data security

II. LITERATURE SURVEY

I. Younus et al. (2015) constructed a new CBIR system that depends on extracting color and texture features. Four extraction methods color moment, color histogram, wavelet moment, and co-occurrence matrices—were used. The authors in Younus et al., (2015) combined k-mean clustering algorithm with particle swarm optimization (PSO), which is a stochastic technique. Tests were conducted by using WANG datasets, which contain 1,000 images divided into 10 classes. Because of incorrect clustering, precision is improved for all classes except architecture and bus as compared to other state-of-the-art techniques. Moreover, the proposed technique did not consider shape feature when computing similarity distance.

II. Anandh et al. (Ponomarev et al., 2016) presented a novel CBIR system based on the integration of color, texture, and shape. Color auto correlogram, Gabor transform, and wavelet transform were used to extract color, shape, and texture, respectively. The authors used Manhattan distance as a similarity measure between the query image and the dataset images. The achieved average precision values were 0.8300, 0.8800, and 0.7000 for the Corel, Li, and Caltech 101 datasets, respectively. The main drawback of the system is the increased computational complexity because of the integration of multiple features. Image analysis at a single resolution level may lose some valuable details.

III. Then, Srivastava and Khare (Srivastava & Khare, (2017) developed a novel multi-resolution analysis algorithm that analyzes images at multiple levels, with other levels capturing information that one level skipped. This approach is based on the extraction of texture and shape features by using the local binary pattern (LBP) descriptor to extract texture features and Legendre moments to extract shape features from the texture features at multi-resolution levels. Although LBP is used to extract local features, it also creates an influential feature vector when local features are combined with global features. Their technique was tested against five datasets, achieving improved accuracy and sensitivity but with increased computational cost due to the use of multi-resolution analysis.

IV. Sajjad et al (2018) proposed an invariant CBIR system to texture rotation and color change. The proposed system based on concatenating color and texture features to form a feature vector with a size of 360. To extract color features, images are converted to HSV color space and quantized through color histogram. To be invariant to illumination change, only Hue and Saturation channels are utilized. Rotated local binary pattern (RLBP) are used to extract rotation invariant texture features. The proposed system is evaluated through experiments on Zurich Building (ZB), Corel 1 K and Corel 10 K.

V. In X. Zheng et al., (2016), authors proposed a CBIR approach based on block processing with overlapping. Firstly, images are transformed to HSI color space then are divided to blocks and the main block is selected. Histogram projection is used to extract color features and Roberts Edge detection is used as texture extraction method. The authors used weighted Euclidean distance as a similarity metric to return similar images and the weights were chosen on experimental basis. The proposed approach has a low accuracy value when compared with other state-of-the-art methods.

VI. A novel CBIR approach is presented by combining color, shape and texture features in Z. Zhao et al., (2016). Color distribution entropy (CDE) was used to extract color features while Hue Moments was used to extract shape features. To extract texture features color level co-occurrence matrix (CLCM) was used. For similarity measurements between query image and dataset images, the authors used weighted normalized similarity measure and the weights were decided upon user's experience. Despite the fact that the proposed system achieves high precision value, the performance of the system is affected when the query image contains more objects (complex). This may be because of using Hue Moments to extract shape features which sometimes does not have the ability to recognize images containing more objects or considers different edges as one edge.

VII. The authors in Phadikar et al., (2018) proposed a CBIR system in compressed domain (Discrete Cosine Domain). Color moments, color histogram and edge histogram have been extracted directly from compressed domain and GA is employed to assign dissimilar importance to the extracted features to improve image retrieval. Although using GA had great positive impact on system's accuracy, it increased the consumption time; however, extracting features in the compressed domain balanced the total time needed to retrieve images.

VIII. A multistage CBIR technique was introduced by Pavithra and Sharmila (Pavithra & Sharmila, (2018). In the first stage, color feature was extracted by using color moment through the calculation of mean and standard deviation for each channel in RGB color space to reduce the search space, which in turn reduce the computational cost. In the second stage, texture and shape (edge) features are extracted from images in the new sub-dataset constructed from the first stage. LBP was used to extract texture information, while Canny edge detector was used to extract edge information. Manhattan distance was used as a metric for the search. Although the proposed multistage system improved performance by increasing precision and decreasing running time, the required running time depends on the number of images in the dataset. If the system is integrated with some machine learning algorithms, then it can be used to search datasets with different sizes and types.

III. EXISTING SYSTEM

The present state of data management systems frequently encounters a number of difficulties and flaws, which calls for the creation of more reliable and secure solutions, such as the project under consideration. Traditional file storage and sharing systems are a common example of an existing system that lacks sophisticated security measures and encryption techniques. These systems usually don't support complex encryption methods or watermarking, and instead rely on simple access restrictions.

DISADVANTAGES:

- Limited Security Measures: Traditional file storage and sharing platforms often lack robust security measures, making them vulnerable to unauthorized access, data breaches, and tampering. Without advanced encryption techniques and watermarking capabilities, the integrity and confidentiality of stored data are compromised.
- Lack of Ownership Verification: In the absence of invisible digital watermarking, traditional systems struggle to provide reliable methods for verifying the ownership and authenticity of uploaded files. This limitation hampers accountability and makes it difficult to track the origins of shared content.
- Vulnerability to Data Theft: Without stringent encryption mechanisms like AES, sensitive information stored within traditional systems is susceptible to interception and theft. Inadequate encryption exposes data to potential breaches during transmission and storage, compromising the privacy and confidentiality of users' files.

IV. PROPOSED SYSTEM

With the introduction of cutting-edge technologies like AES encryption and undetectable digital watermarking, the suggested system seeks to completely transform data management and collaboration. Data owners, users, and administrators may engage with confidence on this powerful and secure platform, which guarantees the integrity, confidentiality, and validity of exchanged information.

ADVANTAGES:

• Strengthened Security Protocols: The suggested solution offers heightened security measures to prevent against unwanted access, data breaches, and manipulation by integrating invisible digital watermarking and AES encryption. While AES encryption protects sensitive data during transmission and storage, watermarking guarantees the legitimacy and ownership of submitted files.

• Trustworthy Ownership Verification: The suggested approach provides a trustworthy way to confirm the ownership and legitimacy of submitted files using invisible digital watermarking. Watermarks inserted into the files provide irrefutable evidence of ownership, which makes data exchanges more accountable and traceable.

• Robust Protection Against Data Theft: By encrypting sensitive data kept in the system, AES encryption provides strong protection against data theft. Even in the case of a breach, the privacy and confidentiality of users' files are protected by this encryption standard, which reduces the possibility of interception and unauthorized access.

.

V. SYSTEM SPECIFICATION

HARDWARE REQUIREMENTS:

- Processor : Intel Core i3 Processor
- Speed : 2.5 GHz
- RAM : 2GB(min)
- Hard Disk : 500MB
- Key Board : Standard Windows Keyboard
- Mouse : Two or Three Button Mouse
- Monitor : LCD

SOFTWARE REQUIREMENTS:

- Operating System : Windows7/10.
- Application Server : Tomcat6.0/7/8.X.
- Front End : Java , HTML,CSS
- Scripts : JavaScript.
- Server side Script : Java Server Pages.
- IDE : Net beans

- Back End : MYSQL 5.0/ Heidi SQL 8.1
- Database Connectivity : JDBC

ARCHITECTURE DIAGRAM:

VI. ENCRYPTION AND WATERMARKING TECHNIQUE

Encryption technique:

We use AES-128 bit to encrypt.-

o Definition: 128-bit encryption is a data/file/key encryption technique that uses a 128-bit key to encrypt and decrypt data or files or keys. 128-bit encryption primarily refers to the length of the encryption or decryption key.

o Code explanation:

In our AES encryption we use this following code format.

1) unsigned char state[] = {0x00, 0x11, 0x22, 0x33, 0x44, 0x55, 0x66, 0x77, 0x88, 0x99, 0xaa, 0xbb, 0xcc, 0xdd, 0xee, 0xff};

2) unsigned char key[] = {0x00, 0x01, 0x02, 0x03, 0x04, 0x05, 0x06, 0x07, 0x08, 0x09, 0x0a, 0x0b, 0x0c, 0x0d, 0x0e, 0x0f};

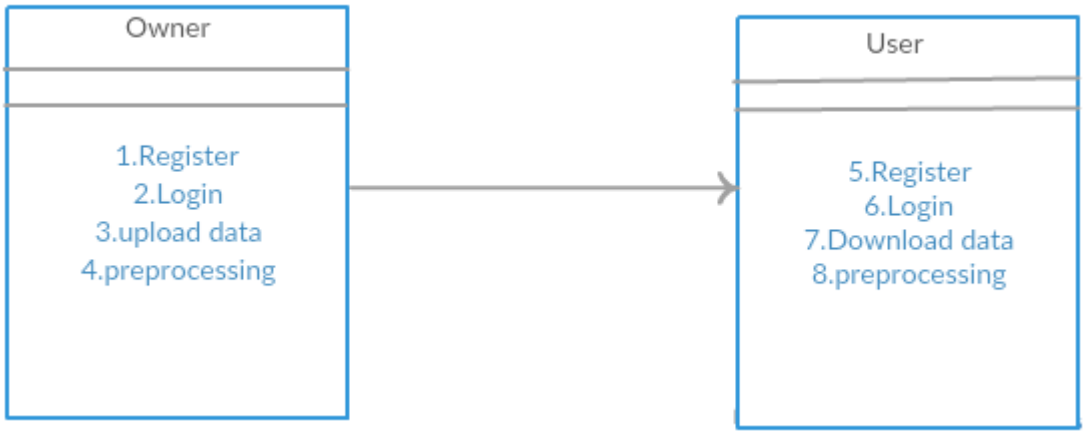
Two arrays of the type unsigned character. Each array is 16 bytes long. The first one contains the plaintext (it means our original data) and the other one the key for the AES encryption (it is an encryption key).

Watermarking technique:

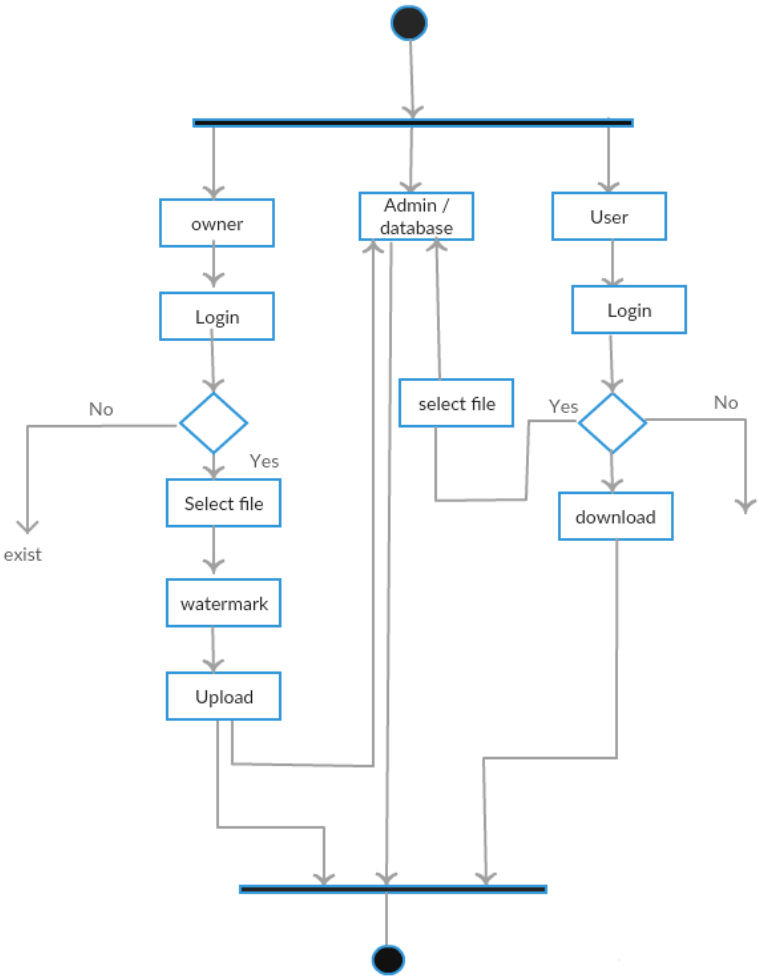
Watermarking embeds a secret message into a cover message. In this project we implement dynamic watermarking. This watermarking technique is used to protect data from attackers. User can't modify the content in the uploaded data's by using watermark. It reduces the attackers.

VII. DIAGRAMS

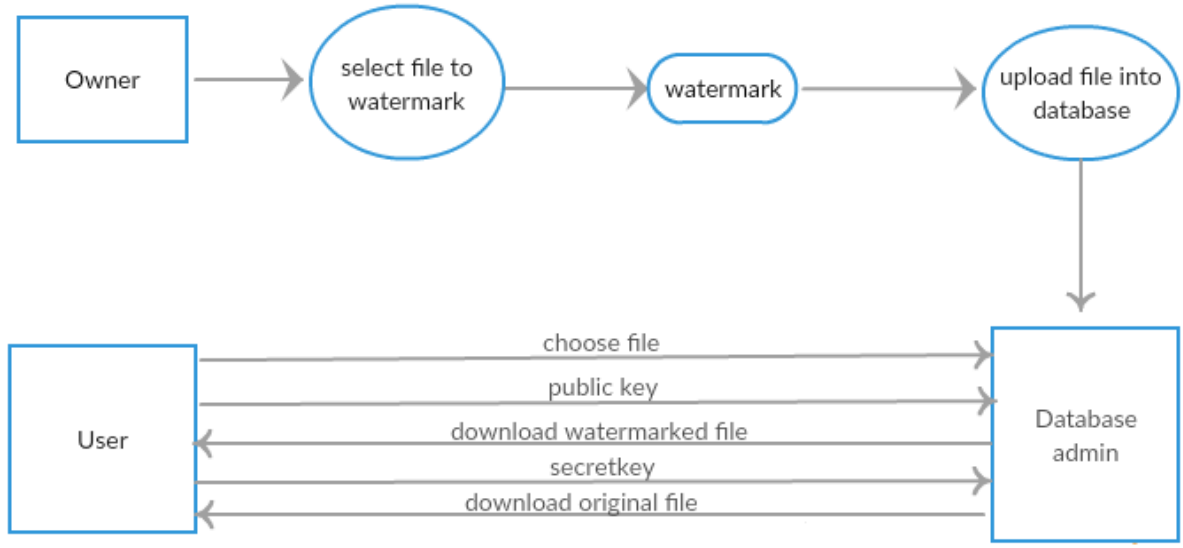
6.1 CLASS DIAGRAM



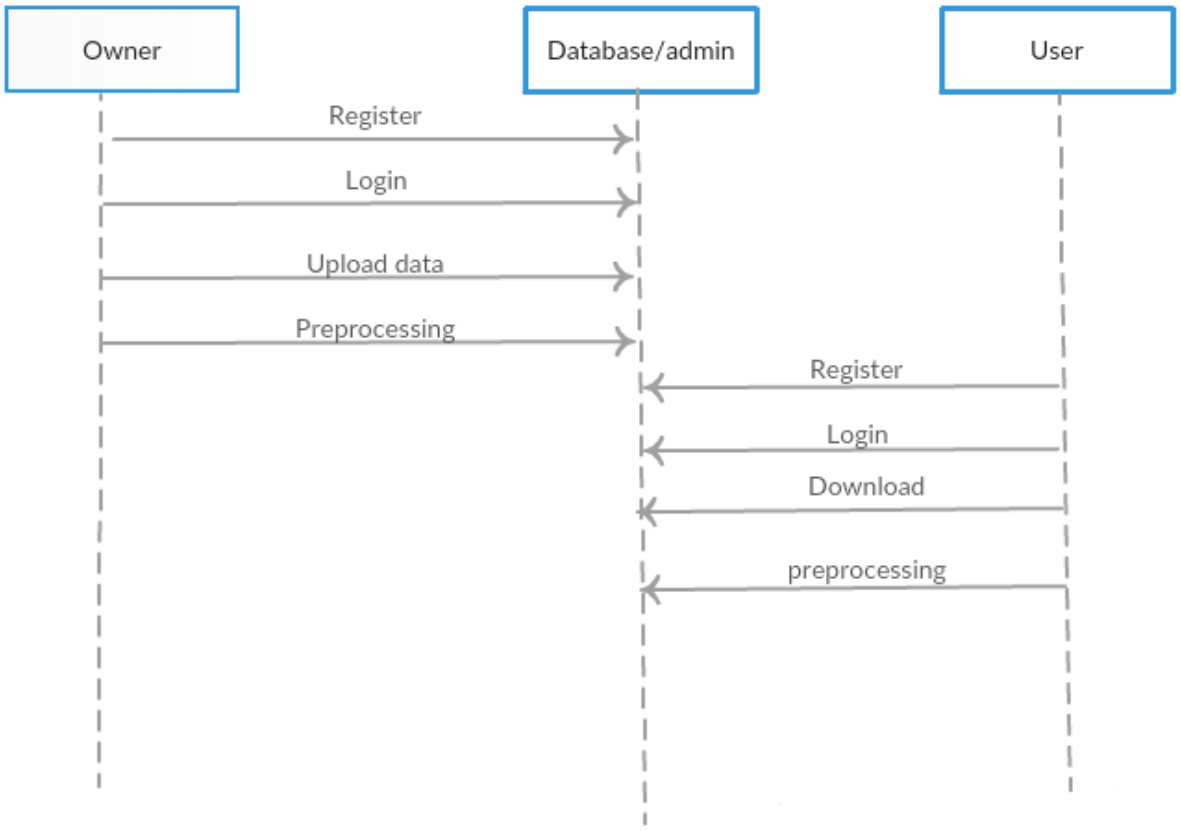
6.2 ACTIVITY DIAGRAM



6.3 DATA FLOW DIAGRAM



6.4 SEQUENCE DIAGRAM



6.5 UML DIAGRAM

VIII. LANGUAGE SPECIFICATIONS:

Java Technology

Java technology is both a programming language and a platform.

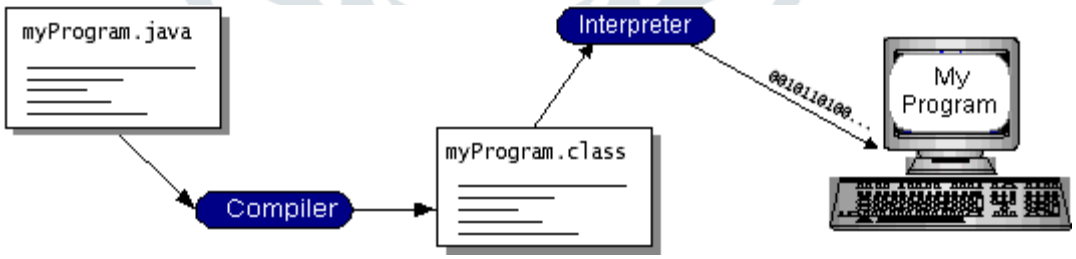
The Java Programming Language

The Java programming language is a high-level language that can be characterized by all of the following buzzwords:

- Simple
- Architecture neutral
- Object oriented
- Portable
- Distributed
- High performance
- Interpreted
- Multithreaded
- Robust
- Dynamic
- Secure

With most programming languages, you either compile or interpret a program so that you can run it on your computer. The Java programming language is unusual in that a program is both compiled and interpreted. With the compiler, first you translate a program into an intermediate language called *Java byte codes* —the platform-independent codes interpreted by the interpreter on the Java platform. The interpreter parses and runs each Java byte code instruction on the computer. Compilation happens just once; interpretation occurs each time the program is executed.

The following figure illustrates how this works.



You can think of Java bytecodes as the machine code instructions for the *Java Virtual Machine* (Java VM). Every Java interpreter, whether it's a development tool or a Web browser that can run applets, is an implementation of the Java VM. Java bytecodes help make "write once, run anywhere" possible. You can compile your program into bytecodes on any platform that has a Java compiler. The bytecodes can then be run on any implementation of the Java VM. That means that as long as a computer has a Java VM, the same program written in the Java programming language can run on Windows 2000, a Solaris workstation, or on an iMac.

IX. SYSTEM TESTING AND MAINTENANCE

The procedure level testing is made first. By giving improper inputs, the errors occurred are noted and eliminated. This is the final step in system life cycle. Here we implement the tested error-free system into real-life environment and make necessary changes, which runs in an online fashion. Here system maintenance is done every months or year based on company policies, and is checked for errors like runtime errors, long run errors and other maintenances like table verification and reports.

UNIT TESTING

Unit testing verification efforts on the smallest unit of software design, module. This is known as "Module Testing". The modules are tested separately. This testing is carried out during programming stage itself. In these testing steps, each module is found to be working satisfactorily as regard to the expected output from the module.

INTEGRATION TESTING

Integration testing is a systematic technique for constructing tests to uncover error associated within the interface. In the project, all the modules are combined and then the entire programmer is tested as a whole. In the integration-testing step, all the error uncovered is corrected for the next testing steps.

X. FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

To guarantee the dependability and applicability of the model in clinical practice, a thorough examination of the outcomes of heart disease prediction using Random Forest should comprise evaluations of accuracy, feature importance, interpretability, validation, and clinical relevance.

XI. CONCLUSION

In conclusion, the project presents a comprehensive solution for secure data management and collaboration in environments. By integrating advanced technologies such as invisible digital watermarking and AES encryption, the proposed system addresses the challenges of ensuring data integrity, authenticity, and confidentiality.The system's ability to embed invisible digital watermarks within shared files enhances ownership verification and traceability, fostering accountability and trust among users. Additionally, AES encryption provides robust protection against unauthorized access, data breaches, and tampering, ensuring the privacy and security of sensitive information.

Through a user-friendly interface and comprehensive administrative tools, the system facilitates seamless interaction and oversight, empowering users and administrators to manage data effectively while enforcing security policies and compliance requirements. Overall, the project represents a significant advancement in data security and management, offering organizations a reliable platform to safeguard their valuable information assets and foster collaboration in a secure digital environment. By

addressing the evolving challenges of data security and compliance, the proposed system lays the groundwork for a more secure and resilient data management infrastructure in the digital age.

XII. REFERENCES

[1] Wang, X., Liu, Y., & Zhang, L. (2023). "A Secure Data Management System with Invisible Digital Watermarking for Cloud-Based Collaborative Environments." Journal of Information Security, 10(3), 127-142.

[2] Chen, H., Li, J., & Xu, Q. (2024). "Enhancing Data Security in Collaborative Systems Using Advanced Encryption and Digital Watermarking." International Journal of Information Security, 18(1), 45-60.

[3] Zhang, H., Liu, W., & Wang, S. (2022). "Invisible Digital Watermarking for Secure File Sharing: A Review of Techniques and Applications." Journal of Cybersecurity and Privacy, 5(2), 88-105.

[4] Smith, A., Johnson, E., & Brown, D. (2023). "Secure Data Management in Cloud-Based Collaborative Environments: Challenges and Solutions." Journal of Computer Security, 15(4), 203-218.

[5] Taylor, L., Wilson, K., & Garcia, R. (2024). "An Integrated Approach to Secure Data Sharing and Collaboration Using Digital Watermarking and Encryption." ACM Transactions on Information and System Security, 27(1), 12-28.

[6] Abdul-Hadi, A. M., Abdulhussain, S. H., Mahmmod, B. M., & Pham, D. (2020, January). On the computational aspects of Charlier polynomials. *Cogent Engineering*, 7(1), 1763553.

[7] Abdulhussain, S. H., & Mahmmod, B. M. (2021, April). Fast and efficient recursive algorithm of Meixner polynomials. *Journal of Real-Time Image Processing*.