Indian Statistical Institute, Kolkata

Project on securing personal health information associated with HIPAA compliance

By

Chandan Ramani

Roll No: CrS1906

under the guidance of Dr. Nagendra Nagaraja

CEO, Founder of QpiAI, Karnataka, Bangalore

Institute supervisor

Dr. Debrup Chakraborty

Head of R C Bose Centre for Cryptology and Security

Indian Statistical Institute, Kolkata

CERTIFICATE

This is to certify that the dissertation entitled as “Securing health information associated with HIPAA compliance” submitted by **Chandan Ramani** to Indian Statistical Institute, Kolkata, in partial fulfilment for the award of the degree of **Master of Technology** in **Cryptology and Security** is a bonafide record of work carried out by him under my supervision and guidance. The dissertation has fulfilled all the requirements as per the regulations of this institute and in my opinion, has reached the standard needed for submission.

Dr. Nagendra Nagaraja

CEO, Founder QpiAI Pvt. Ltd.

Bangalore, 560045

Date : July , 2021

Acknowledgement

I would first like to express my profound gratitude to my advisor, Dr. Nagendra Nagaraja, CEO, Founder, QpiAI Pvt. Ltd. ,Bangalore, India, for his patient guidance, enthusiastic encouragement, and useful critiques needed for this research work. His advice and knowledge helped me in pursuing good research and writing of this thesis. I am privileged to have such a great supervisor as he consistently supported me both academically and personally.

My grateful thanks are also extended to my institute supervisor Dr. Debrup Chakraborty, R C Bose Centre for Cryptology and Security, Indian Statistical Institute, Kolkata, India, for his valuable advice and assistance in keeping my progress on schedule. I am privileged to have such a expert as a guide.

I want to thank Ms. Swati Kumari, Director of Management, QpiAI Pvt. Ltd for guiding me throughout this internship.

Finally, I want to thank my parents, family, friends for making everything easier and smooth to do this project.

Chandan Ramani
M.Tech CrS
Indian Statistical Institute
Kolkata - 700108, India.

Abstract

In this project my aim was to research about how to secure personal health information associated with HIPAA compliance. Firstly, we tried to understand what is HIPAA and how it is connected to health information. Then we tried to find related encryption schemes for health data. Also, it was needed to find authentication schemes in this project. Our goal was to find how to save health data securely in cloud system and how to access that data with proper authentication. All this work was done keeping in mind that everything should be HIPAA compliant.

Contents

page no

1. Introduction

Security and Privacy are very essential in the healthcare industry. It should be known that with whom data can be shared and what data should be shared. Patient's trust and carrying privacy is the foundation of a success of a healthcare system. A patient may not willingly disclose of their health information in the absence of a trusted environment and this could have bad effects. So there should be a strict guideline and protocol for that. Country wise there are different acts and rules. Here we will deal with HIPAA.

1.1 What is HIPAA ?

HIPAA stands for Health Insurance Portability and Accountability Act (1996), America. A federal law, restrictive or permissible law to protect health information. It covers information that created, received or maintained on behalf of health care providers and health plans. It was thought to improve health efficiency and effectiveness of health care system. Also, it was realized that, as digitalization has been started so we need some administrative steps to protect health information.

HIPAA covered entities are Health providers -Doctors, Health plans -Insurers, Health care clearing houses, Business associates.
Not HIPAA covered entities are …Data on our phone -like shopping history, Data on fitness tracker app, Data related to Tech Companies, any data related to any person until that person have not shared his/her data with any covered entities.

1.2 What are HIPAA associate rules?

There are lots of rules regarding this. In our work we followed mainly privacy and security related rules. Here some of them are pointed.

Privacy: The privacy rule defines authentication in **§164.312(d)** as
"Implement procedures to verify that a person or entity seeking access to electronic protected health information is the one claimed."

HIPAA sets limits and conditions on the uses and disclosures that may be made of Protected health info without patient's authorization. HIPAA gives patients right over their health information to examine, to copy records and to request corrections.

Security: The covered entity must **164.312(e)(2)(ii)**

"Implement a mechanism to encrypt electronic protected health information whenever deemed appropriate."

HIPAA establishes set of security standards to protect Personal health information (PHI). It provides technical/non-technical safeguards that covered entities (Providers – Doctors,_Health plans – Insurers, Health care clearinghouses, Business Associates) must put in place to secure individuals electronic personal health information(ePHI). When Business Associate (BA) connect with covered entity, privacy rule requires that covered entity include certain protections for the information.

HIPAA Enforcement rule (2009) : To strengthen civil and criminal enforcement of the HIPAA rules and to handle significantly increased civil monetary penalties for violations. In this scenario Health and Human Services (HHS) has responsibilities for HIPAA violations.

Access Control rule : The Security Rule defines access in **§164.304** as *"the ability or the means necessary to read, write, modify, or communicate data/information or otherwise use any system resource."*

Audit Control rule : The Security Rule defines audit in **§164.312(b)** as *"Implement hardware, software, and/or procedural mechanisms that record and examine activity in information systems that contain or use electronic protected health information."*

1.3 What can we disclose under HIPAA?

a) Treatment related data – referral, consultations
b) Payment related data – premium, reimbursement
c) Health care operations – legal, administrative, staff evaluations
d) Victims of abuse, domestic violence
e) Cadaveric organ donations
f) Govt. functions
g) Research
h) Worker's compensation
i) Serious threats

1.4 Challenges in Security and Privacy of health data

a) Verification of user

b) Integrity and availability

c) Access control

d) Data control

d) Profiles of users

e) Misuse of health records

2. Information Security

Information security includes three topics. Access control, secure communication and protection of the private data.

Access control: Includes both the initial entrance by a participant and the re-entry of that participant and the access of additional participants i.e., who can rightfully access computer system or data.

An example of access control at the application level is the setting of cookies by browser.

Secure Communication: This includes any transfer of information among any of participants. The most commonly recognized function of a secure system is the encryption algorithm and most commonly recognized problem in a secure system is the encryption key management. Some of the other functions and issues for security systems are **hashing**(for checking data integrity), **identity authentication** (for allowing access), **electronic signatures** (for preventing revocation of legitimate transactions), **information labelling**(for tracing location and times for transactions), and **monitors**(for identifying potential attacks on the system).

Protection of the private data: This includes storage devices, processing units, cache memory etc.

The level of security required is not universal. Ease of access is more important for low-security activities, such as reading advertisements. More difficult access is required for medium security such as bank accounts. High security is required for high-value corporate proprietary computations, such as design data for a next-generation product.

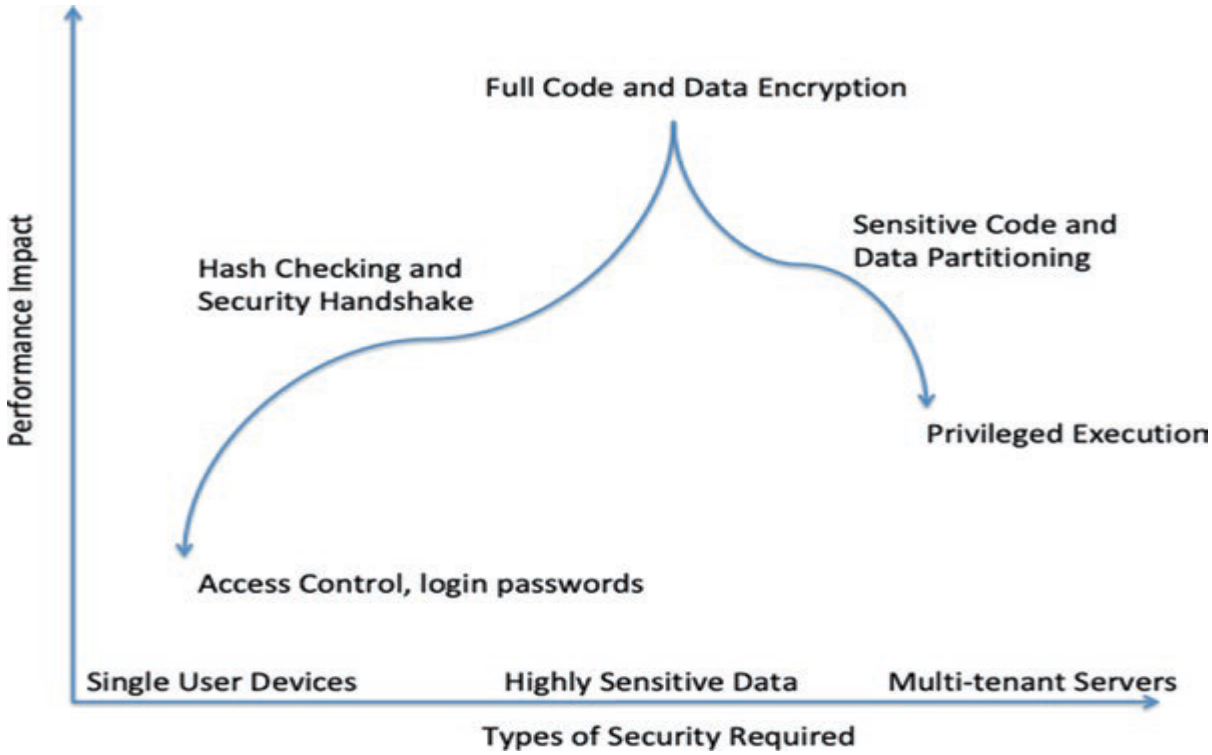

Fig 1 Performance vs Security

3. Security Concerns of Cloud Operating Models

Software as a Service (SaaS): Highest layer of abstraction focused on end-users of Cloud, to provide them application access, such that multiple users can share the same application binary in their own virtual machines or server instances. An application provider wants to ensure that users have a read-only access to the binary and all patches are updated in a timely manner.

Platform as a Service (PaaS): Focused on application developers, providing them access to elastic servers that can stretch in CPU cores, memory, and storage on need basis. The PaaS requires strong middleware security. PaaS permits integrations of services, so the security perimeters need a greater degree of access control.

Infrastructure as a Service (IaaS): The bottommost layer in a Cloud stack, providing direct access to virtualized or containerized hardware. Users want to ensure that hardware level services, such as various drivers and ports, are protected from other processes running on the same physical server.

4. Useful encryption schemes

There are various encryption schemes used to encrypt health information and authentication techniques. Advantages and disadvantages in each encryption techniques and authentication techniques are also discussed in this section.

4.1 Private/Symmetric key encryption:

A symmetric-key encryption took the form of a map $E: K \times M \rightarrow C$, so that the map $E_k: M \rightarrow C$, $m \rightarrow E(k, m)$ is referred to as encryption function for each $k \in K$.

$D_k = E^{-1}(k)$ is referred as the decryption function.

Symmetric key encryption uses a large key size, it is safe against attack by brute force. Moreover, in both hardware and software, it can be implemented efficiently because it generates strong keys.

Applicability of symmetric key encryption in encrypting health information is not convenient. Because the secret key should be shared between the doctor and the patient. So if a doctor has many patients, the doctor has to maintain those many keys for each patient. Also a doctor has to communicate with other doctors regarding patient conditions, with the pharmacy regarding drugs and many more, so each party in the whole system has to maintain many numbers of keys which is very difficult.

4.2 Public/Asymmetric key encryption:

Public key encryption scheme is based on the hardness of mathematical problem. There is a possibility of man in the middle attack in Public key encryption. The adversary may modify the sender's messages by pretending that the adversary is the original sender to the receiver. Having a trusted third party called certificate authority to issue the certificate to the public key can solve this man in the middle attack. The sender receives the receiver's public key and certificate and verifies the public key. If the verification process succeeds the sender use receivers public key to encrypt the messages.

The major drawback of certificate authority Public key encryption is to maintain the certificates which are complex and cumbersome.

4.3 Identity based encryption:

In 1984, Shamir introduced identity-based encryption to overcome the certificate management in public key encryption. The user's public key is an arbitrary string that is a user identifier rather than a number. A trusted third party called a private key generator, generates the users' private key. Both the sender and the receiver

must authenticate to private key generator in order to obtain their private key that corresponds to their identity.

The main drawback is, it suffers key escrow problem and the communication between the private key generator and the users should be secure to transfer private key.

4.4 Attribute based encryption:

Few years later, Fuzzy Identity-Based Encryption is proposed, which is also known as Attribute-Based Encryption (ABE). In such encryption scheme, an identity is viewed as a set of descriptive attributes, and decryption is possible if a decrypter's identity has some overlaps with the one specified in the ciphertext. More general tree-based ABE schemes, Key-Policy Attribute-Based Encryption (KP-ABE) and Ciphertext-Policy Attribute-Based Encryption (CP-ABE), are presented to express more general condition than simple overlap.

ABE suffers from inflexibility due to complex access control policies. In addition, the In this scheme, the data owner has to maintain all the attributes that are burdensome in a single authority.

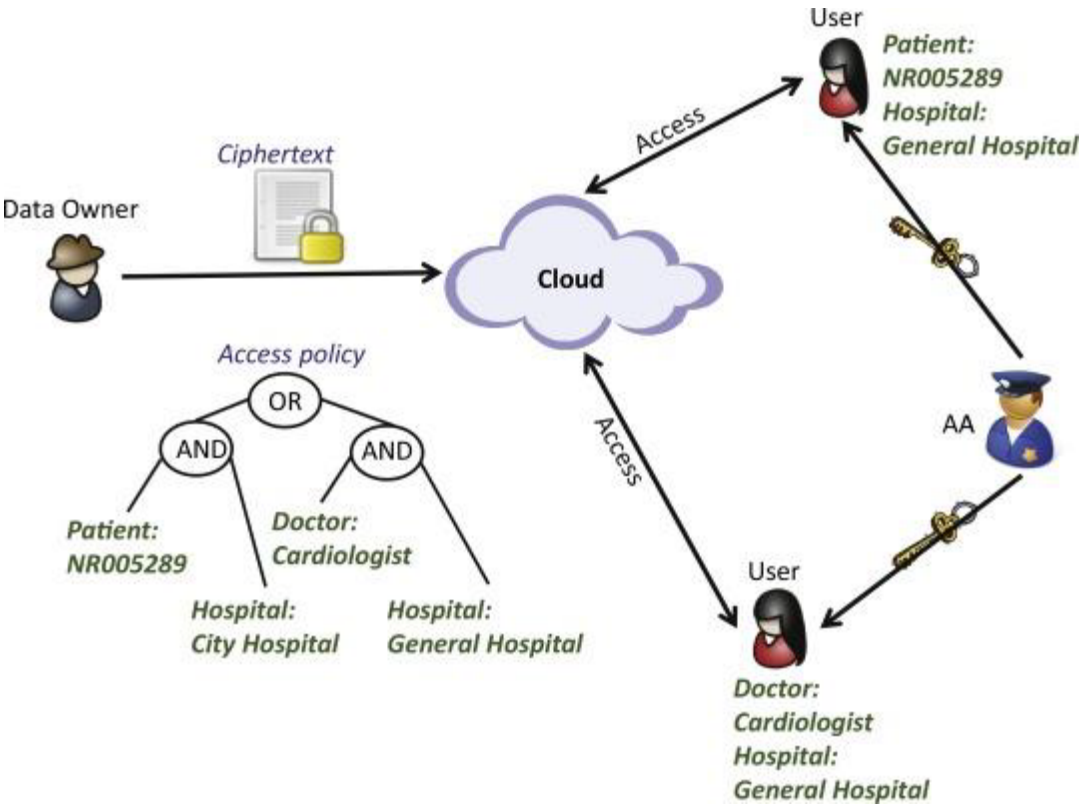

Fig 2, structure of the system

5. Authentication Schemes

Authentication mean data authentication and user authentication. Authentication of data ensures the source of the data. Authentication of the user ensure the authenticity of the user. There are various authentication schemes as follows.

5.1 Digital signature scheme: Digital signature schemes are used for data authentication. User authentication process consists of two keys which are bounds to the user's identity. Certificates for the user's key has to be maintained, which is cumbersome and costly. So certificate less digital signature schemes are developed by adopting hierarchical identity based digital signature schemes. Comparing with all digital signature schemes, biometric-based signature schemes provide stronger security by avoiding man in the middle attack.

5.2 Username/Password: Username/password mechanism used to secure health data by user authentication. This mechanism does not guarantee data integrity as a man in the middle attack is susceptible.

5.3 Login/password with digital certificate: HIPAA regulation advice two-factor authentication. As for example, login contains the username or email address anything identifying the user. Along with it, a digital certificate is needed for the user to prove his authenticity. Biometrics of user can be used instead of the digital certificate to improve the security.

5.4 Credential systems: In this scheme, a trusted third party will issue credentials to the users. The user holding legitimate credential can access the data outsourced to the cloud. A better solution relative to previous ones.

5.5 Biometric based system: Biometric-based authentication provides stronger security for data which are outsourced to the cloud. It avoids man in the middle attack.

Recent research says biometric can be reframed from the details. So employing biometric template are risky.

5.6 Block Chain Technology: Definition of Block Chain is as follows.

"Block Chain is an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way" - lansiti, Lakhani.

Here, open = accessible to all, distributed = not centralized, efficiently = fast and scalable, verifiable = everyone can check the validity, permanent = persistent of information.

Block Chain contains two parts -one is header part, another is transaction/data part. Header of a block connects transactions-any change in transaction will change in block header. If someone want to change something in data part, whole chain will be updated as it uses Merkle Tree structure containing Hash values.

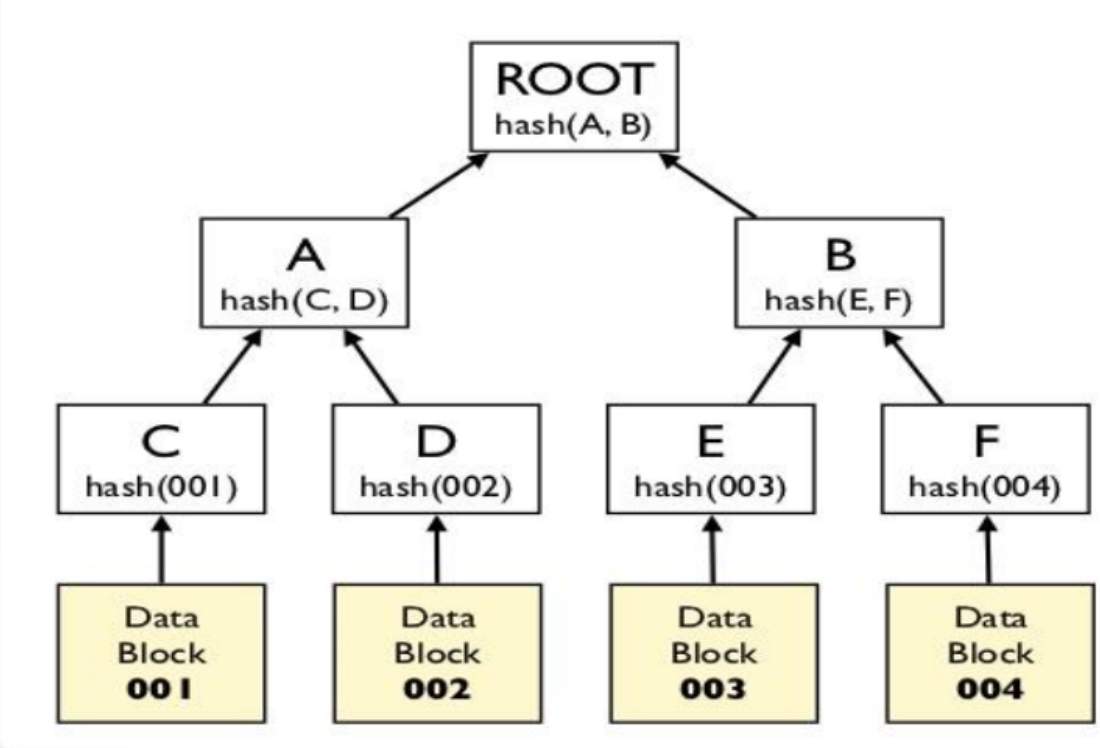

Fig 3, Merkle Tree/ Hash Tree

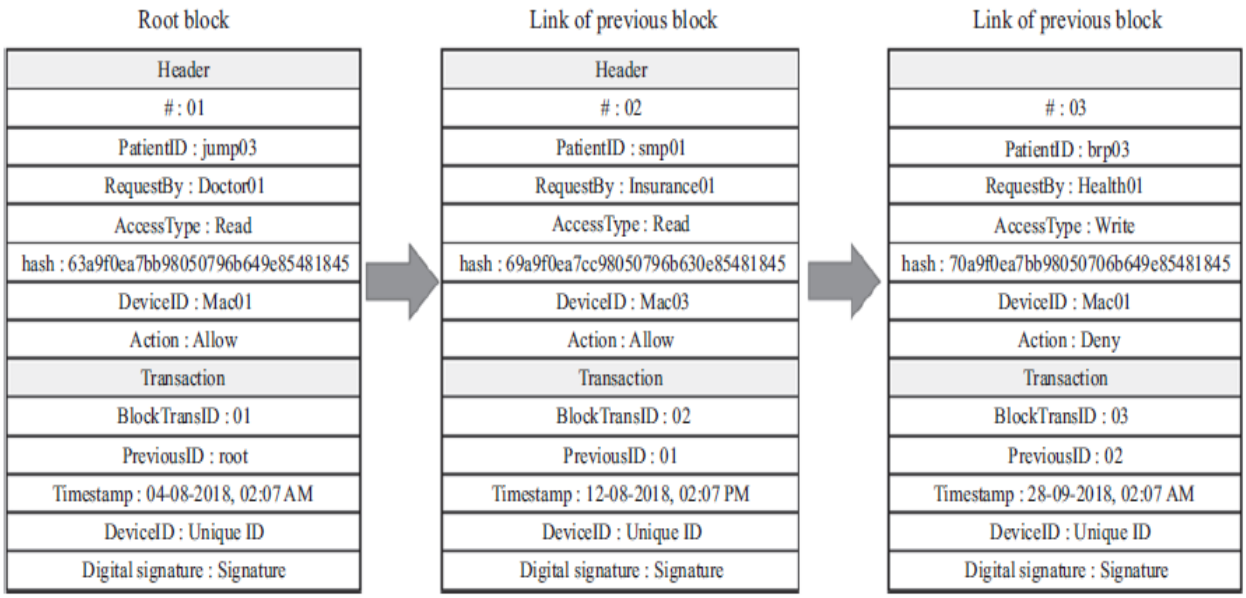

Fig 4, Sample of Blocks for health data in Block Chain model

Security of Block Chains are built on the concept that Hash Functions are cryptographically secure.

6. Future Work

Here through out this chapters, I have mentioned about current structures like, encryption schemes, authentication schemes. But that can be modify more. For encryption multiauthority attribute-based encryption, Hierarchical attribute-based encryption, all these can be followed. For authentication and storing we can more rely on Block Chain technology, though it has some drawbacks.

Research is going on rapidly regarding security and privacy of health data. Researchers are trying to find solutions to make this system more patient centric. Cloud technology has some major drawbacks like maintenance issues, hardware set up, data fragmentation, high cost etc.

Centralized server may have advantages, but is has lots of danger, if something wrong happens then all data will be disclosed or damaged.

More awareness needed. Patient should have knowledge on data management system. Health care personals should be trained more.

7. Conclusion

There is no perfect solution for solving the security issues. Nothing is 100% secure. What we can do is preventing attacks by following some protocols. HIPAA, HL7 etc standards should be followed while designing and developing software for security and privacy. Training to staff and physicians, about proper data entry, modification and data safeguarding should be imparted. Some of protocols are as follows:

- All the hardware equipment, namely, server, storage, networking equipment like a switch, access point, security equipment like UTM, Firewall, etc. should have appropriate physical security. There should be an access control system (based on PIN, card, biometric or combination of them) for accessing the systems.
- All the hardware should be properly maintained by time. While communicating the data, secure communication modes like https, vsftpd, scp, etc. can be used. Network segmentation is another way to prevent unauthorized access up to a certain extent.
- There should be proper backup and restore mechanism.
- All the software components (application or system) should be regularly patched and updated or upgraded whichever possible. Frequent security audit and mock drills must be conducted to safeguard the system from cybersecurity threats.
- In the software, proper identity management for authentication and role management for authorization should be present. Alerts should be generated in case of any unauthorized access. Block chain technology can be adopted for authentication and access control.
- Operating system hardening should be done on the servers and end-user systems from where the application is accessed.
- All the data should be encrypted while storing. Full disk encryption may also be used.

8. References

1. https://www.hhs.gov/sites/default/files/nist-csf-to-hipaa-security-rule-crosswalk-02-22-2016-final.pdf
2. https://www.hhs.gov/sites/default/files/cyber-attack-checklist-06-2017.pdf
3. https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/securityrulepdf.pdf?language=es
4. https://www.researchgate.net/publication/11007067_Implementing_security_and_access_control_mechanisms_for_an_electronic_healthcare_record
5. Control Cloud Data Access Privilege and Anonymity With Fully Anonymous Attribute-Based Encryption by
   Taeho Jung, Xiang-Yang Li, *Senior Member, IEEE*, Zhiguo Wan, and Meng Wan, *Member, IEEE*
6. Cloud Computing with Security by
   Naresh Kumar Sehgal, Pramod Chandra P. Bhatt, John M. Acken
7. Cloud Computing Security by John R. Vacca
8. techsafeguards.pdf
9. Analysis of Various Encryption Algorithms in Cloud Computing by
   Nasarul Islam.K.V, Mohamed Riyas.K.V

HELLOSIGN

Audit trail

TITLE	Final Report_Chandan
FILE NAME	Final_Report_Chandan.docx
DOCUMENT ID	e82b02e3cf5dca10d82f7551d0bf138b5a680ccd
AUDIT TRAIL DATE FORMAT	MM / DD / YYYY
STATUS	Completed

Document history

SENT	**07 / 08 / 2021** 06:53:01 UTC	Sent for signature to Nagendra Nagaraj (nagendra.nagaraja@qpiai.tech) from payroll.qpiai@qpiai.tech IP: 183.82.155.106
VIEWED	**07 / 08 / 2021** 07:37:07 UTC	Viewed by Nagendra Nagaraj (nagendra.nagaraja@qpiai.tech) IP: 122.166.226.159
SIGNED	**07 / 08 / 2021** 07:38:15 UTC	Signed by Nagendra Nagaraj (nagendra.nagaraja@qpiai.tech) IP: 122.166.226.159
COMPLETED	**07 / 08 / 2021** 07:38:15 UTC	The document has been completed.