ISSN: 2583-9888
DOI: doi.org/10.51219/JAIMLD/ankit-srivastava/374

URF PUBLISHERS
connect with research world

Journal of Artificial Intelligence, Machine Learning and Data Science

https://urfpublishers.com/journal/artificial-intelligence

Vol: 2 & Iss: 4

Research Article

Securing PHI Data In Healthcare

Ankit Srivastava*

Citation: Srivastava A. Securing PHI Data In Healthcare. J Artif Intell *Mach Learn & Data Sci 2024*, 2(4), 1678-1679. DOI: doi.org/10.51219/JAIMLD/ankit-srivastava/374

Received: 03 October, 2024; **Accepted:** 28 October, 2024; **Published:** 30 October, 2024

***Corresponding author:** Ankit Srivastava, USA, E-mail: ankit1985sri@gmail.com

ABSTRACT

PHI data privacy has increasingly become a matter of concern in the era of large public digital repositories of data. This is particularly true in healthcare where data can be misused if traced back to patients, and brings with itself a myriad of possibilities. Digital data in healthcare is a double-edged sword. While on the one hand, digitalization has allowed for a wide variety of advancements, including teleconsultations, easy retrieval and duplication of data for records and development of applications such as machine learning, it has also allowed for the possibility that the personal medical records of a patient can be accessed by a number of individuals.

Keyword: Healthcare, PHI, HIPPA, Cloud, Data Security, ETL

1. Introduction

"PHI data" in healthcare stands for "Protected Health Information," which refers to any individually identifiable health information about a patient, including demographic details, medical history, test results, and insurance information, that is held by a covered entity like a healthcare provider or health plan and must be protected under the Health Insurance Portability and Accountability Act (HIPAA) regulations; essentially, any information that could be used to identify a patient and is related to their past, present, or future health condition or healthcare treatment.

2. What is PHI?

Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means. Further, PII is defined as information: (i) that directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address, etc.) or (ii) by which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator and other descriptors). Additionally, information permitting the physical or online contacting of a specific individual is the same as personally identifiable information. This information can be maintained in either paper, electronic or other media[1].

Personally Identifiable Information (PII) is any information that can be used to identify an individual.

PHI can include a person's name, address, phone number, email address, Social Security number, driver's license number, passport number, or financial account number, person's face, fingerprints or handwriting.

PHI can also include information that indirectly identifies a person, such as their gender, race, birth date or geographic indicator.

PHI also include information about a person's medical history, mental or physical condition, or medical treatment, health insurance policy number or subscriber identification number.

3. HIPPA Security rules

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) required the Secretary of the U.S. Department of Health and Human Services (HHS) to develop regulations

protecting the privacy and security of certain health information Securing PHI data in cloud. The Security Rule applies to health plans, health care clearinghouses, and to any health care provider who transmits health information in electronic form in connection with a transaction for which the Secretary of HHS has adopted standards under HIPAA (the "covered entities") and to their business associates[2]. There are numerous regulatory compliance guidelines that carry penalties in the event of a PHI breach. The largest regulatory framework covering PHI is HIPAA. According to the U.S. Department of Health and Human Services (HHS), the HIPAA Privacy Rule "provides federal protections for personal health information held by covered entities and gives patients an array of rights with respect to that information."

The Security Rule requires covered entities to maintain reasonable and appropriate administrative, technical, and physical safeguards for protecting e-PHI.

- Specifically, covered entities must:
- Ensure the confidentiality, integrity, and availability of all e-PHI they create, receive, maintain or transmit.
- Identify and protect against reasonably anticipated threats to the security or integrity of the information.
- Protect against reasonably anticipated, impermissible uses or disclosures; and

Ensure compliance by their workforce

4. Securing PHI Data in Cloud

Now a days more companies are transitioning from on premise data base to cloud and securing phi data in cloud is challenge since companies are using service as SaaS bases. There are multiple ways to secure the data in cloud on way is to store the PHI information in encryption form. Data can be stored in 64-byte encryption or 256-byte encryption. Another way to secure data in cloud is restricting the access of PHI columns to user, give access of PHI columns access to selected users only. To secure data from cyber attack there should be multi-level of authentication and authorization, and anti-virus should be thoroughly updated. Employees should also be properly trained in identifying any phishing email and securing data in case of potential cyber-attack. HIPAA requires that organizations collect and analyze audit logs related to PHI access to detect suspicious activity. It should go without saying that organizations should regularly review audit logs to make sure they remain in compliance[3]. The loss of PII can result in substantial harm to individuals, including identity theft or other fraudulent use of the information. Because DOL employees and contractors may have access to personal identifiable information concerning individuals and other sensitive data, we have a special responsibility to protect that information from loss and misuse.

5. Securing PHI data during ETL load

ETL process involved transforming the data from one system to another. extracting the data from source system doing transformation and loading the data in target. to protect data while transforming there should be Continuously monitor data access activities and maintain detailed logs to identify any suspicious behavior or potential security breaches. For PHI data transformation secured connections can be created in source and target. data should be encrypted on fly for extra security layer. There should be extra security in the target table where PHI data is loaded. ETL security is not an option; it's a necessity in today's data-driven world. Failing to secure your ETL processes can lead to data breaches, regulatory non-compliance, and reputational damage. By following the best practices outlined above, you can protect your data at every stage of the ETL pipeline and ensure that it remains secure and confidential[5].

6. AI Based Data Protection

The "black box" nature of AI systems means their decision-making processes often have opacity. This obscurity raises concerns for businesses, users, and regulators, as they often cannot see or understand how AI algorithms arrive at certain conclusions or actions. A lack of algorithmic transparency can also obscure biases or flaws in AI systems, leading to outcomes that may inadvertently harm certain groups or individuals. Without this transparency, businesses risk eroding customer confidence and potentially breaching regulatory requirements.

Since data transfer can result in leaks of data and is particularly problematic in case of transfer across the border, attempts have been made at transferring networks, rather than data. Federated learning is a sort of distributed learning in which several clients work together to jointly develop a model, while maintaining the confidentiality of their input. Here the learning happens separately, each time with a separate set of data, and the model trained ultimately can draw from knowledge across all datasets[4].

7. Secure data in Cyber attack

To secure PII (Personally Identifiable Information) data in a cyber-attack, key strategies include: identifying and classifying PII, implementing strict access controls, encrypting data both at rest and in transit, regularly monitoring for suspicious activity, educating employees on data protection practices and having a robust incident response plan in place to minimize damage in case of a breach; essentially, limiting access to sensitive data, making it unreadable without authorization, and proactively preparing for potential attacks.

8. Conclusion

In conclusion, protecting PHI (Protected Health Information) through robust data security practices is crucial for healthcare organizations to maintain patient trust, comply with HIPAA regulations, and mitigate the significant risks associated with data breaches, which can lead to legal repercussions, financial losses, and reputational damage for individuals and organizations alike; therefore, implementing comprehensive administrative, physical and technical safeguards to

9. References

1. https://www.dol.gov/general/ppii
2. https://www.hhs.gov/hipaa Pub. L. 104-191
3. Charles Wang, Lead Product Evangelist, Fivetran.
4. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 2019;10:1-19.
5. 2024 Gartner® Magic Quadrant™ for Endpoint Protection Platforms.